

Db 203 GPDBRLYTFCKSLPTKFSKMLDLNCT-SSVPVVKVMVEECNCE 248

RESULT 14

O35793 PRELIMINARY; PRT; 184 AA.

AC O35793;

DT 01-JAN-1998 (TrEMBLrel. 05, Created)

DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE DR protein.

GN DR.

OC Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=97378055; PubMed=9234736;

RA Topol L.Z., Marx M., Laugier D., Bogdanova N.N., Boubnov N.V.,

RA Clausen P.A., Calothy G., Blair D.G.;

RT "Identification of drm, a novel gene whose expression is suppressed in

RT transformed cells and which can inhibit growth of normal but not

RT transformed cells in culture.";

RL Mol. Cell. Biol. 17:4801-4810(1997).

DR EMBL; Y10019; CAA71126.1; -.

DR InterPro; IPR000359; Cys_knot.

DR Pfam; PF03045; DAN_dom.

DR SMART; SM00041; DAN; 1.

SQ SEQUENCE 184 AA; 20680 MW; 90AC99749B1ADB8F6 CRC64;

Query Match 8.3%; Score 86; DB 11; Length 184;

Best Local Similarity 23.3%; Pred. No. 0.46;

Matches 37; Conservative 24; Mismatches 60; Indels 38; Gaps 9;

QY 4 QAFRNDATVPIGLGEYPEPPENNQTWNAENGCRPPHPHYDAKDVSYSCRELHYT-- 61

Db 38 KAQNDSEQT-----QSPPOGSRTRGRGQ--GRGTAMP--GEEVLESSQEALHVT 86

QY 62 RFLTDGPCR-----SAKPVTELVCSGCGPARLLPNAIGRVKWRPNPDPFR 108

Db 87 KYLRDWCKTQPLKQTIHEEGNSRTIINRFYGCQN-SFYIPRH-----RKEGSGFQ 139

QY 109 ----CIPDRYRAQRVQLLCPGGAAP-RSRKRVLVASCKC 142

Db 140 SCSFCKPKKFTTMMVTLNCPPELPPTKKKRVTRVKQCRC 178

RESULT 15

O70326 PRELIMINARY; PRT; 184 AA.

AC O70326;

DT 01-AUG-1998 (TrEMBLrel. 07, Created)

DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE GREMLIN (Cysteine KNOT superfamily 1, BMP antagonist 1).

GN CTSF1B1 OR DRM.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98325381; PubMed=9660951;

RA Hsu D.R., Economides A.N., Wang X., Eimon P.M., Harland R.M.;

RT "The Xenopus dorsaling factor Gremlin identifies a novel family of

RT secreted proteins that antagonize BMP activities.";

RL Mol. Cell 1:673-683(1998).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE-SPLEEN;

RA Zhang Q.Y., Topol L.Z., Athanasiou M., Blair D.G.;

"Cloning of the murine DRM gene and characterization of its oncogene suppressible promoter.";

RT Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.

RL [3]

RP SEQUENCE FROM N.A.

RC TISSUE=COLON;

RA Strausberg R.;

RL EMBL; AF045801; AAC40111.1; -.

DR EMBL; AF108189; AAD54056.1; -.

DR EMBL; BC015293; AAH15293.1; -.

DR MGD; MGI:1344337; Cktsf1b1.

DR InterPro; IPR000359; Cys_knot.

DR Pfam; PF03045; DAN_dom.

DR SMART; SM00041; CT; 1.

SQ SEQUENCE 184 AA; 20710 MW; D1BD99783CDC0F2D CRC64;

Query Match 8.3%; Score 86; DB 11; Length 184;

Best Local Similarity 23.3%; Pred. No. 0.46;

Matches 37; Conservative 24; Mismatches 60; Indels 38; Gaps 9;

QY 4 QAFRNDATVPIGLGEYPEPPENNQTWNAENGCRPPHPHYDAKDVSYSCRELHYT-- 61

Db 38 KAQNDSEQT-----QSPPOGSRTRGRGQ--GRGTAMP--GEEVLESSQEALHVT 86

QY 62 RFLTDGPCR-----SAKPVTELVCSGCGPARLLPNAIGRVKWRPNPDPFR 108

Db 87 KYLRDWCKTQPLKQTIHEEGNSRTIINRFYGCQN-SFYIPRH-----RKEGSGFQ 139

QY 109 ----CIPDRYRAQRVQLLCPGGAAP-RSRKRVLVASCKC 142

Db 140 SCSFCKPKKFTTMMVTLNCPPELPPTKKKRVTRVKQCRC 178

Search completed: March 23, 2003, 14:18:47

Job time : 54.28 secs

Mon Mar 31 09:44:37 2003

```

RL Cytogenet. Cell Genet. 89:79-84 (2000).
DR EMBL; AF045800; AAC39725.1; -.
DR EMBL; AB032372; BAA84462.1; -.
DR EMBL; AF110137; AAF06677.1; -.
DR EMBL; AF154054; AAG23891.1; -.
DR InterPro; IPR000359; Cys knot.
DR InterPro; IPR004133; DAN_dom.
DR Pfam; PF03045; DAN; 1.
DR SMART; SM00041; CT; 1.
DR SEQUENCE 184 AA; 20697 MW; 4B588598DE12C47E CRC64;

Query Match      8.6%; Score 88.5; DB 4; Length 184;
Best Local Similarity 23.4%; Pred. No. 0.25;
Matches 36; Conservative 20; Mismatches 55; Indels 43; Gaps 8;

Qy 23 PPNENQ-----TNRAENGGRPPHPHYDAKDVSEYSCRELHYT--RFLTD 66
Db 34 PPPDKACHNDSEQTSPQPGSRNRGRGQGTAMP--GEVLESSQEALHVTTERKYLK 91

Qy 67 GPCR-----SAKPVTELVCSGCGPARLLPNAIGRVKWRPNPNDPFR---C 109
Db 92 DWCKTQPLKQTHIEBGCNSRTIINRFYVGCN-SFYIPRHI-----RKEEGSFQSCSFC 144

Qy 110 IPDRYRAORVQLCPGGAP-RSRKVRIVASCKC 142
Db 145 KPKKFTTMMVTLNCPPELOPPTKKKRVTRVKQCR 178

RESULT 11
Q8WNY1 PRELIMINARY; PRT; 184 AA.
AC Q8WNY1;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Gremlin.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
SEQUENCE FROM N.A.
RA Christenson L.K., Duffy D.M.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF439783; AAL32022.1; -.
DR InterPro; IPR000359; Cys knot.
DR InterPro; IPR004133; DAN_dom.
DR Pfam; PF03045; DAN; 1.
DR SMART; SM00041; CT; 1.
DR SEQUENCE 184 AA; 20697 MW; 4B588598DE12C47E CRC64;

Query Match      8.6%; Score 88.5; DB 6; Length 184;
Best Local Similarity 23.4%; Pred. No. 0.25;
Matches 36; Conservative 20; Mismatches 55; Indels 43; Gaps 8;

Qy 23 PPNENQ-----TNRAENGGRPPHPHYDAKDVSEYSCRELHYT--RFLTD 66
Db 34 PPPDKACHNDSEQTSPQPGSRNRGRGQGTAMP--GEVLESSQEALHVTTERKYLK 91

Qy 67 GPCR-----SAKPVTELVCSGCGPARLLPNAIGRVKWRPNPNDPFR---C 109
Db 92 DWCKTQPLKQTHIEBGCNSRTIINRFYVGCN-SFYIPRHI-----RKEEGSFQSCSFC 144

Qy 110 IPDRYRAORVQLCPGGAP-RSRKVRIVASCKC 142
Db 145 KPKKFTTMMVTLNCPPELOPPTKKKRVTRVKQCR 178

RESULT 12
Q40600 PRELIMINARY; PRT; 332 AA.
ID Q40600
AC Q40600;

```

```

DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE 70S mitochondrial ribosomal protein L2.
GN RP12.
OS Oenothera lutea (Bertero's evening primrose).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Myrtales; Onagraceae; Oenothera.
OX NCBI_TaxID=3950;
RN [1]
SEQUENCE FROM N.A.
RA Fritz I., Schuster W.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; X80170; CAA56451.1; -.
DR InterPro; IPR002171; Ribosomal_L2.
DR Pfam; PF00181; Ribosomal_L2; 1.
DR Ribosomal protein.
KW Ribosome.
SQ SEQUENCE 332 AA; 36678 MW; B8E123914F01791B CRC64;

Query Match      8.5%; Score 88; DB 10; Length 332;
Best Local Similarity 25.5%; Pred. No. 0.54;
Matches 27; Conservative 16; Mismatches 35; Indels 28; Gaps 5;

Qy 3 WOAFNRDATEVIGLGEYPPPPENNTNNRAENGGRPPHPHYDAKDVSEYSCRELHYTR 62
Db 244 WLSFRQETDGLVGAAEHNSKPKTDQ-----GSLP-----AKPIGEGT----- 282

Qy 63 FLTDPGCR-SAKPVTELVCSGCGPARLLPNAIGRVKWRPNPNDP 107
Db 283 --KDGACKVDRAPTYIIASHOLEAGKWMVNC---DWSKPSTSD 322

RESULT 13
Q9PWB0 PRELIMINARY; PRT; 272 AA.
AC Q9PWB0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Cerberus homology.
GN CER.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
SEQUENCE FROM N.A.
RA MEDLINE=99439862; PubMed=10508582;
RA Zhu L., Marvin M.J., Gardiner A., Lassar A.B., Mercola M., Stern C.D.,
RA Levin M.;
RA "Cerberus regulates left-right asymmetry of the embryonic head and
RT heart."
RL Curr. Biol. 9:931-938 (1999).
DR EMBL; AF139721; AAD51610.1; -.
DR InterPro; IPR000359; Cys knot.
DR InterPro; IPR004133; DAN_dom.
DR Pfam; PF03045; DAN; 1.
DR SMART; SM00041; CT; 1.
DR PROSITE; PS01225; CTCK_2; 1.
DR SEQUENCE 272 AA; 31224 MW; 36E4C9F719711BCA CRC64;

Query Match      8.4%; Score 87; DB 13; Length 272;
Best Local Similarity 21.5%; Pred. No. 0.55;
Matches 23; Conservative 20; Mismatches 42; Indels 22; Gaps 4;

Qy 44 PYDAKDVSEYSCRELHYTRFLTPGCRSAKPVTELVCSGCGPARLLPNAIGRVKWRPN 103
Db 157 PIKTNMHQETCTRLPFSQVAHESCEKV-IVQNNLCFGKCS-----FHPV 202

Qy 104 GPDR-----CIPDRYRAORVQLCPGGAPRSRKRVRIVASCKCK 143

```


DR InterPro: IPR000359; Cys knot.
 DR InterPro: IPR004133; DAN_dom.
 DR Pfam: PF03045; DAN; 1.
 DR SMART: SM00041; CT; 1.
 SQ SEQUENCE 184 AA; 21166 MW; 3510B44E8FDD5EA CRC64;

Query Match 8.8%; Score 90.5; DB 13; Length 184;
 Best Local Similarity 22.9%; Pred. No. 0.16;
 Matches 35; Conservative 23; Mismatches 60; Indels 35; Gaps 8;

QY 18 GEYEP-----PPNNQTMNRAENGRRPHI-----PYDAKDVSEYSCRELHYT--RFLTDG 67
 DB 33 GAIPPPKQDPNDSEQMTQQSSRRHREGKGTSPABEVLESSQEAUHYTERKYLKRD 92
 QY 68 PCR-----SAKPVTELVCSCGCGPARLLP-----NAIGRVKWRPNPGRPPFR-----CI 110
 DB 93 WCKTQPLKQTHGEGCNRITINFCYGCN-SFYIPRHV-----RKEGSGFQSCSFCK 145

QY 111 PDYRAQRVQLLCFPGAAPRSK-VRLVASC 142
 DB 146 PKFTTMTVTLNCPQLPQPRKKKTRVKECRC 178

RESULT 5
 QY557 ID QY557 PRELIMINARY; PRT; 712 AA.
 AC QY557;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE DJ1163J1.3 (Similar to mouse B99 protein) (Fragment).
 GN DJ1163J1.3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lloyd D.;
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL031588; CAB38415.1; -.
 FT NON TER 1
 SQ SEQUENCE 712 AA; 75685 MW; ECBD2133EF72F32 CRC64;

Query Match 8.7%; Score 90; DB 4; Length 712;
 Best Local Similarity 22.7%; Pred. No. 0.78;
 Matches 34; Conservative 20; Mismatches 36; Indels 60; Gaps 6;

QY 68 PCSAKPVTELVCSCGCGPARLLP-----NAIGRVKWRPNPGRPPFCIPDRYARV----- 119
 DB 333 PANSSRPLSNISKSGRMGPAMLRLPALPAGPVGASSW-----QAKRVDVSE 377

QY 120 ----QLLCPGGAAP-----RSRKVLVASC-----KCKELT 146
 DB 378 LAAEQLTAPPASPTQPTPEGGGWLNSSCAWSESSQLNKTSTRIRRDSCLSNKTWKMP 437

QY 147 RFHQSELKDF-----GPETARPQKGRKPRP 172
 DB 438 TPTNQFKIPKFSIGDSPDSTPKLSRAQRP 467

RESULT 6
 QY523 ID QY523 PRELIMINARY; PRT; 720 AA.
 AC QY523;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE B99.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;

[1] SEQUENCE FROM N.A.
 RP MEDLINE=20432105; PubMed=10974554;
 RX Monte M., Collavin L., Lazarevic D., Utrera R., Dragani T.A.,
 RA Schneider C.;
 RT "Cloning, chromosome mapping and functional characterization of a
 human homologue of murine Gtse-1 (B99) gene.";
 RL Gene 254:229-236(2000).
 DR EMBL; AF223408; AAF31459.1; -. D4E852CE4BE85BBF CRC64;
 SQ SEQUENCE 720 AA; 76628 MW; 76628 MW; D4E852CE4BE85BBF CRC64;

Query Match 8.7%; Score 90; DB 4; Length 720;
 Best Local Similarity 22.7%; Pred. No. 0.79;
 Matches 34; Conservative 20; Mismatches 36; Indels 60; Gaps 6;

QY 68 PCSAKPVTELVCSCGCGPARLLP-----NAIGRVKWRPNPGRPPFCIPDRYARV----- 119
 DB 341 PANSSRPLSNISKSGRMGPAMLRLPALPAGPVGASSW-----QAKRVDVSE 385

QY 120 ----QLLCPGGAAP-----RSRKVLVASC-----KCKELT 146
 DB 386 LAAEQLTAPPASPTQPTPEGGGWLNSSCAWSESSQLNKTSTRIRRDSCLSNKTWKMP 445

QY 147 RFHQSELKDF-----GPETARPQKGRKPRP 172
 DB 446 TPTNQFKIPKFSIGDSPDSTPKLSRAQRP 475

RESULT 7
 QY580 ID QY580 PRELIMINARY; PRT; 720 AA.
 AC QY580;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE G-2 and S-phase expressed 1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Straussberg R.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC006325; AA06325.1; -.
 SQ SEQUENCE 720 AA; 76614 MW; ACD91CCCD008A89C CRC64;

Query Match 8.7%; Score 90; DB 4; Length 720;
 Best Local Similarity 22.7%; Pred. No. 0.79;
 Matches 34; Conservative 20; Mismatches 36; Indels 60; Gaps 6;

QY 68 PCSAKPVTELVCSCGCGPARLLP-----NAIGRVKWRPNPGRPPFCIPDRYARV----- 119
 DB 341 PANSSRPLSNISKSGRMGPAMLRLPALPAGPVGASSW-----QAKRVDVSE 385

QY 120 ----QLLCPGGAAP-----RSRKVLVASC-----KCKELT 146
 DB 386 LAAEQLTAPPASPTQPTPEGGGWLNSSCAWSESSQLNKTSTRIRRDSCLSNKTWKMP 445

QY 147 RFHQSELKDF-----GPETARPQKGRKPRP 172
 DB 446 TPTNQFKIPKFSIGDSPDSTPKLSRAQRP 475

RESULT 8
 QY786 ID QY786 PRELIMINARY; PRT; 988 AA.
 AC QY786;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Gastric mucin (fragment).

RC	TISSUE=KIDNEY;
RA	Strausberg R.;
RL	Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
DR	EMBL; AK007967; BAB25378.1; -
DR	EMBL; AK002240; BAB21957.1; -
DR	EMBL; AK002396; BAB22068.1; -
DR	EMBL; AK007893; BAB25333.1; -
DR	EMBL; BC021458; AAB21458.1; -
DR	MGD; MGI:1913292; O61006G05Rik.
DR	InterPro; IPR00359; Cys_knot.
DR	PROSITE; PS01225; CTCK_2; 1.
SQ	SEQUENCE 206 AA; 23174 MW; 70D24819EB06CBC9 CRC64;
Query Match 32.0%; Score 330; DB 11; Length 206;	
Best Local Similarity 39.8%; Pred. No. 2.8e-26;	
Matches 74; Conservative 35; Mismatches 65; Indels 12; Gaps 7;	
QY	5 AFRRDATEVIPGLGEYPPEP-PENNOTMNRANGRRPPHPH---YDAKDVSSEYSCRELHY 60
DB	23 AFKNDATEILYSHVVKVPAPHPSSNSTLNQANGGR--HFSSTGLDNRSRVQVGCRELS 80
QY	61 TRFLTDGPCRSKAKPVTELVCSQCQPALLPNAIG---RVKMW-RPNGDPDFRCIPDRYRA 116
DB	81 TKYSISDGOCTISPLKELVCAGECLPLPVLPNWIGGGYGTYKWSRRSSQEWRCVNDKTRT 140
QY	117 QRVOLLCPGGAAPRSRRVLRVLVASCKCKELTRFNQSELKDFGPETARP-QXGRKRPFGAK 175
DB	141 QRILQOCODGST-RYKITYVTACKRYTRQHNESSHFNSFSPAKPAQHHRKRKASK 199
QY	176 ANQAEI 181
DB	200 SSKHSL 205
RESULT 2	
Q96HU7	PRELIMINARY; PRT; 206 AA.
ID	Q96HU7
AC	Q96HU7;
DT	01-DEC-2001 (TrEMBLrel. 19, Created)
DI	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE	Similar to RIKEN CDNA 0610006G05 gene (CDA019).
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_Taxid=9606;	[1]
RN	SEQUENCE FROM N.A.
RP	TISSUE=BONE MARROW;
RC	Strausberg R.;
RA	Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
RL	EMBL; BC008484; AAF08484.1; -
DR	EMBL; AF361494; AAL57219.1; -
DR	InterPro; IPR00359; Cys_knot.
DR	PROSITE; PS01225; CTCK_2; 1.
SQ	SEQUENCE 206 AA; 23306 MW; 9FB3CC41E4B53834 CRC64;
Query Match 31.8%; Score 328; DB 4; Length 206;	
Best Local Similarity 39.9%; Pred. No. 4.5e-26;	
Matches 73; Conservative 36; Mismatches 62; Indels 12; Gaps 7;	
QY	5 AFRRDATEVIPGLGEYPPEP-PENNOTMNRANGRRPPHPHPIADKDV---EYSCELHY 60
DB	23 AFKNDATEILYSHVVKVPAPHPSSNSTLNQANGGR--HFSYTGLDNRTRVQVGCRELS 80
QY	61 TRFLTDGPCRSKAKPVTELVCSQCQPALLPNAIG---RVKMW-RPNGDPDFRCIPDRYRA 116
DB	81 TKYSISDGOCTISPLKELVCAGECLPLPVLPNWIGGGYGTYKWSRRSSQEWRCVNDKTRT 140

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:12:16 ; Search time 53.28 Seconds

(without alignments)
715.441 Million cell updates/sec

Title: US-09-867-274-4

Perfect score: 1032

Sequence: 1 QGWAQFRDATEVIGLGEY.....KGRKPRPGAKANQALEWAY 185

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 21:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_rvirus:*

16: sp_bacteriap:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	330	32.0	206	11 Q9CQK4	Q9CQK4 mus musculus
2	328	31.8	206	4 Q96HJ7	Q96HJ7 homo sapien
3	249	24.1	134	4 Q9X3U3	Q9X3U3 homo sapien
4	90.5	8.8	184	13 Q73755	Q73755 gallus gall
5	90	8.7	712	4 Q9Y557	Q9Y557 homo sapien
6	90	8.7	720	4 Q9NYZ3	Q9NYZ3 homo sapien
7	90	8.7	720	4 Q9BRE0	Q9BRE0 homo sapien
8	90	8.7	988	6 Q97867	Q97867 sus scrofa
9	88.5	8.6	182	13 Q73754	Q73754 xenopus lae
10	88.5	8.6	184	4 Q60565	Q60565 homo sapien
11	88.5	8.6	184	6 Q8WNY1	Q8WNY1 macaca mula
12	88	8.5	332	10 Q40600	Q40600 cenothera b
13	87	8.4	272	13 Q9PWB0	Q9PWB0 gallus gall
14	86	8.3	184	11 Q35793	Q35793 rattus norv
15	86	8.3	184	11 Q70326	Q70326 mus musculus
16	86	8.3	1081	4 Q76065	Q76065 homo sapien

17	85.5	8.3	215	8 Q9MFB6	Q9MFB6 beta vulgar
18	85	8.2	272	13 Q9PUK2	Q9PUK2 gallus gall
19	85	8.2	1035	5 Q9NEG1	Q9NEG1 drosophila
20	85	8.2	1114	11 Q9JKW7	Q9JKW7 mus musculus
21	84.5	8.2	1223	11 Q9QW33	Q9QW33 rattus sp.
22	83.5	8.1	279	10 Q8VZ42	Q8VZ42 arabidopsis
23	83	8.0	644	12 Q65732	Q65732 bluetongue
24	83	8.0	644	12 Q65751	Q65751 bluetongue
25	83	8.0	667	10 Q8S1G5	Q8S1G5 oryza sativ
26	82.5	8.0	141	6 Q9S1J85	Q9S1J85 monodelphis
27	82.5	8.0	168	11 Q88273	Q88273 mus musculus
28	82.5	8.0	334	10 Q49558	Q49558 arabidopsis
29	82	7.9	1637	6 Q9XSV8	Q9XSV8 bos taurus
30	82	7.9	5146	6 Q8SPM4	Q8SPM4 bos taurus
31	81.5	7.9	262	13 Q13091	Q13091 pleurodeles
32	81.5	7.9	392	10 Q9AX94	Q9AX94 oryza sativ
33	80.5	7.8	157	6 Q8WNC7	Q8WNC7 pongo pygma
34	80.5	7.8	157	6 Q8WNC5	Q8WNC5 pongo pygma
35	80.5	7.8	168	4 Q9H772	Q9H772 pongo pygma
36	80.5	7.8	270	13 Q70041	Q70041 homo sapien
37	80.5	7.8	386	10 Q8S5P2	P70041 xenopus lae
38	80.5	7.8	1704	5 Q94446	Q8S5P2 oryza sativ
39	80.5	7.8	384	11 Q9CUP6	Q94446 chironomus
40	80	7.8	766	13 Q98SW2	Q9CUP6 mus musculus
41	80	7.8	1042	4 Q13792	Q98SW2 oncorhynchus
42	80	7.8	157	6 Q8WNC6	Q13792 homo sapien
43	79.5	7.7	157	6 Q8WNC6	Q8WNC6 pongo pygma
44	79	7.7	141	6 Q8WN18	Q8WN18 alluropoda
45	79	7.7	2026	4 Q00468	Q00468 homo sapien

ALIGNMENTS

RESULT 1

ID Q9CQK4 PRELIMINARY; PRT; 206 AA.

AC Q9CQK4;
DT 01-JUN-2001 (TREMREL. 17, Created)
DT 01-JUN-2001 (TREMREL. 17, Last sequence update)
DT 01-MAR-2002 (TREMREL. 20, Last annotation update)
DE 0610006G05RIK protein (RIKEN CDNA 0610006G05 gene).
GN 0610006G05RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CE7BL/6J; TISSUE=PANCREAS, AND KIDNEY;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Mateu Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stauble P., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohseki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.

Search completed: March 28, 2003, 14:16:52
Job time : 16.3067 secs

RA McNab D., Perry L.J., Scott J.E., Taylor P.;
RT "The complete DNA sequence of the long unique region in the genome of
RL herpes simplex virus type 1";
CC J. Gen. Virol. 69:1531-1574 (1988).
CC -!- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL49,
CC EHV-1 11, AND VZV 9.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; D10879; BAA01695.1; -
DR EMBL; X14112; CAA32259.1; -
DR PIR; D30089; WMBEF9
SQ SEQUENCE 301 AA; 32254 MW; 6595539C2ABE13E29 CRC64;

Query Match 7.9%; Score 82; DB 1; Length 301;
Best Local Similarity 26.1%; Pred. No. 2;
Matches 43; Conservative 17; Mismatches 67; Indels 38; Gaps 10;

QY 21 PEPENNO---TMRANGRPHPHYDAKDVSEY-----SCRELHYTRFLTDGCRSAK 73
DB 33 PDPSPDTSRRGALQTRSQRGEVRFVQYDESDYALGGSSSEDEH-----PEVP-RTRR 86
QY 74 PVELVCSQCQCPARLLP-----NAIGRVKWRNPGDFRCIPRYAQRVOLLCPGAA 128
DB 87 PVSGAVLSGP-GPARAPPPGAGGAGTPTTAPRAP-----RTORVATKAP--AA 134
QY 129 PRSRKRVLVASCKRLTRFNQSELKDFGPETARPOKGRKPRG 173
DB 135 PAATTR-----GRKSAQPEAAALPD-APASTAPTRSKTPAQG 171

RESULT 14
CABI_RAT
ID CABI_RAT STANDARD; PRT; 2182 AA.
AC O88480;
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Calcineurin-binding protein Cabin 1 (Calcineurin inhibitor) (CAIN).
GN CABIN1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hippocampus;
RX MEDLINE=98325042; PubMed=9660798;
RA Lai M.M., Burnett P.E., Wolosker H., Blackshaw S., Snyder S.H.;
RT "Cain, a novel physiological protein inhibitor of calcineurin.";
RL J. Biol. Chem. 273:18325-18331(1998).
CC -!- FUNCTION: INTERACTS WITH AND INHIBITS CALCINEURIN-MEDIATED SIGNAL
CC TRANSDUCTION. MAY PROVIDE A DOCKING SITE FOR CALCINEURIN IN ITS
CC INACTIVE FORM. MAY TARGET INACTIVATED CALCINEURIN TO SPECIFIC
CC INTRACELLULAR LOCATIONS WHERE ITS RELEASE WOULD PROVIDE CALCIUM-
CC REGULATED PHOSPHATASE ACTIVITY TO SPECIFIC SIGNALING PATHWAYS.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: HAS A PROMINENT NEURONAL EXPRESSION AND A WIDE
CC TISSUE DISTRIBUTION.
CC -!- SIMILARITY: CONTAINS 6 TPR REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AF061947; AAC40176.1; -
DR InterPro; IPR001440; TPR.
DR Pfam; PF00515; TPR; 4.
DR SMART; SM00028; TPR; 3.
KW Repeat; TPR repeat.
FT REPEAT 36 69 TPR 1.
FT REPEAT 90 123 TPR 2.
FT REPEAT 125 157 TPR 3.
FT REPEAT 615 648 TPR 4.
FT REPEAT 1055 1088 TPR 5.
FT REPEAT 1106 1139 TPR 6.
FT DOMAIN 2078 2115 CALCINEURIN BINDING.
SQ SEQUENCE 2182 AA; 242811 MW; 2904C335AB3F440A CRC64;

Query Match 7.9%; Score 82; DB 1; Length 2182;
Best Local Similarity 22.8%; Pred. No. 15;
Matches 50; Conservative 17; Mismatches 68; Indels 84; Gaps 9;

QY 4 QAFRNDATFVIP-----GLGEYPEPP-----PENNOTMNAENG-----R 39
DB 1693 KASPEDGQESLPHPKKLPADGSGPGPEGKVGKGLHQLPVATDTRDNTQSGGPKDKR 1752
QY 40 PPHHPYDAKDVSEYSCR--ELHYT-RLTDPGCRSAKP-----VTELVCSCQCGPAR 88
DB 1753 PVGPTFPMDTGETAARHPDLEPTPRLPGPRDRGPFESRSAELELSISTRQOPAP 1812
QY 89 LLPNAIGRVKWRNPGDFRCIPRYAQRVOLLCPGGAAPRSRKRVLVASCKRLTRF 148
DB 1813 LVPSPV-----TPTAAPTNGARAAGH----- 1835
QY 149 HNQSELKDFGPETA--RPQGRKPRGAKANQAELENAY 185
DB 1836 -----PEAPPRPNRRKRLQDTESGKTLILDAY 1864

RESULT 15
LSHB_CERSI
ID LSHB_CERSI STANDARD; PRT; 141 AA.
AC O77835; O19102;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lutropin beta chain precursor (luteinizing hormone beta subunit) (LSH-
DE beta) (LSH-B) (LH-B).
GN LHB1 AND LHB2.
OS Ceratotherium simum (White rhinoceros) (Square-lipped rhinoceros).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Ferissodactyla; Rhinocerotidae; Ceratotherium.
OX NCBI_TaxID=9807;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98389253; PubMed=9723860;
RA Lund L.A., Sherman G.B.;
RT "Duplication of the southern white rhinoceros (Ceratotherium simum
RT simum) luteinizing hormone beta subunit gene.";
RL J. Mol. Endocrinol. 21:19-30(1998).
RN [2]
RP SEQUENCE OF 7-141 FROM N.A.
RC TISSUE=pituitary;
RX MEDLINE=97449288; PubMed=9305757;
RA Sherman G.B., Lund L.A., Bunick D., Winn R.J.;
RT "Characterization and phylogenetic significance of rhinoceros
RT luteinizing hormone beta (LHbeta) subunit messenger RNA structure,
RT complementary DNA sequence and gene copy number.";
RL Gene 195:131-139(1997).
CC -!- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING
CC THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.
CC -!- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
CC CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC -!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN

Mon Mar 31 09:44:37 2003

```

FT DISULFID 114 121 BY SIMILARITY.
FT CARBOHYD 34 34 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 141 AA; 15060 MW; C6CF98036B3C4EE0 CRC64;

Query Match 8.4%; Score 86.5; DB 1; Length 141;
Best Local Similarity 28.8%; Pred. No. 0.36; Indels 17; Gaps 5;
Matches 34; Conservative 9; Mismatches 59;

QY 55 CRELHYTRFLTDGCRSAKPVTELVCSGQC-GPARLLPNAIGRVKWRPNPGDPFCIPDR 113
DB 30 CRPTNATLAESDACPVCVFTTICAGYCPSMVRVLPAL-----PPGQLVCTYRE 82
QY 114 YRAQRVOLL-CPGAAP-RSRKRVLVASCKKELTRFHNSQLKDFGPETARPQKGRFP 170
DB 83 LSFSSIRLPGCPGVDPIFSPFVALSCSGCSKLSH-----SDCGGRAPRPHICTRP 134

RESULT 11
VP4_BTIV11 STANDARD; PRT; 644 AA.
AC P33428;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE VP4 core protein.
GN S4.
OS Bluetongue virus (serotype 11 / isolate USA).
OC Viruses; dsRNA viruses; Reoviridae; Orbivirus.
OX NCBI_TaxID=33716;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93331736; PubMed=8393245;
RA Huang I.J., Hayama E., Jeong Y.J., Li J.K.-K.;
RT "Conservation of the segment 4 gene sequence and of a leucine zipper motif in VP4 among five US bluetongue viruses.";
RL Virology 195:772-779(1993).
CC -!- FUNCTION: THE VP4 PROTEIN IS ONE OF THE FIVE PROTEINS (WITH VP1, VP3, VP6 AND VP7) WHICH FORM THE INNER CAPSID OF THE VIRUS.
CC -!- SIMILARITY: BELONGS TO THE REOVIRUSES VP4 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L08638; AAA42825.1; -.
CC Core protein.
KW SEQUENCE 644 AA; 75443 MW; 578503CFCD9B37D6 CRC64;
SQ

Query Match 8.1%; Score 84; DB 1; Length 644;
Best Local Similarity 22.8%; Pred. No. 2.8;
Matches 53; Conservative 16; Mismatches 73; Indels 90; Gaps 9;

QY 8 NDATVIFGLGEYEPPEPPNNQTNRAENGRRPHHPYDAKVS-----EYSCELHYTR 62
DB 141 NDANPRHLKVIYGMPEIP---LYMEYAEIGARFDDEPTDEKLVSLMDYIVYSAEHHVY- 196
QY 63 FLTDGCRSAKPVTELVCSGQCQPALL-----PNAIGRVKWRPNPGDPFCIPDR 113
DB 197 -----GCGDLRTLMQFKRSRFRVRLVHWYDPIAPE--CSDPN 234
QY 114 YRAQRVOLLCPGAAPRSRKRVLVASCKKELTRF-----HNQSELKDFGPETARPQK 167
DB 235 VIVHNIWV-----DSKKILKMNFLKRVRLFTWVSDRSQMDHEWETTRFAED 286
QY 168 RK-----PRPGAKANQAELEN 183
DB 287 RLGEIAYEMGGAFSSALIKHRIPNSKDEYHCISTYLFQPGADADMYELRN 338

Query Match 8.1%; Score 84; DB 1; Length 644;
Best Local Similarity 22.8%; Pred. No. 2.8;
Matches 53; Conservative 16; Mismatches 73; Indels 90; Gaps 9;

QY 8 NDATVIFGLGEYEPPEPPNNQTNRAENGRRPHHPYDAKVS-----EYSCELHYTR 62
DB 141 NDANPRHLKVIYGMPEIP---LYMEYAEIGARFDDEPTDEKLVSLMDYIVYSAEHHVY- 196
QY 63 FLTDGCRSAKPVTELVCSGQCQPALL-----PNAIGRVKWRPNPGDPFCIPDR 113
DB 197 -----GCGDLRTLMQFKRSRFRVRLVHWYDPIAPE--CSDPN 234
QY 114 YRAQRVOLLCPGAAPRSRKRVLVASCKKELTRF-----HNQSELKDFGPETARPQK 167
DB 235 VIVHNIWV-----DSKKILKMNFLKRVRLFTWVSDRSQMDHEWETTRFAED 286
QY 168 RK-----PRPGAKANQAELEN 183
DB 287 RLGEIAYEMGGAFSSALIKHRIPNSKDEYHCISTYLFQPGADADMYELRN 338

```

```

RESULT 12
VP4_BTIV11 STANDARD; PRT; 644 AA.
AC P33429;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE VP4 core protein.
GN S4.
OS Bluetongue virus (serotype 13 / isolate USA).
OC Viruses; dsRNA viruses; Reoviridae; Orbivirus.
OX NCBI_TaxID=33717;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93331736; PubMed=8393245;
RA Huang I.J., Hayama E., Jeong Y.J., Li J.K.-K.;
RT "Conservation of the segment 4 gene sequence and of a leucine zipper motif in VP4 among five US bluetongue viruses.";
RL Virology 195:772-779(1993).
CC -!- FUNCTION: THE VP4 PROTEIN IS ONE OF THE FIVE PROTEINS (WITH VP1, VP3, VP6 AND VP7) WHICH FORM THE INNER CAPSID OF THE VIRUS.
CC -!- SIMILARITY: BELONGS TO THE REOVIRUSES VP4 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L08640; AAA42827.1; -.
CC Core protein.
KW SEQUENCE 644 AA; 75268 MW; BA2FEF769BA490E0 CRC64;
SQ

Query Match 8.1%; Score 84; DB 1; Length 644;
Best Local Similarity 22.8%; Pred. No. 2.8;
Matches 53; Conservative 16; Mismatches 73; Indels 90; Gaps 9;

QY 8 NDATVIFGLGEYEPPEPPNNQTNRAENGRRPHHPYDAKVS-----EYSCELHYTR 62
DB 141 NDANPRHLKVIYGMPEIP---LYMEYAEIGARFDDEPTDEKLVSLMDYIVYSAEHHVY- 196
QY 63 FLTDGCRSAKPVTELVCSGQCQPALL-----PNAIGRVKWRPNPGDPFCIPDR 113
DB 197 -----GCGDLRTLMQFKRSRFRVRLVHWYDPIAPE--CSDPN 234
QY 114 YRAQRVOLLCPGAAPRSRKRVLVASCKKELTRF-----HNQSELKDFGPETARPQK 167
DB 235 VIVHNIWV-----DSKKILKMNFLKRVRLFTWVSDRSQMDHEWETTRFAED 286
QY 168 RK-----PRPGAKANQAELEN 183
DB 287 RLGEIAYEMGGAFSSALIKHRIPNSKDEYHCISTYLFQPGADADMYELRN 338

RESULT 13
UL49_HSV11 STANDARD; PRT; 301 AA.
AC P10233;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE Tegument protein UL49.
GN UL49.
OS Herpes simplex virus (type 1 / strain 17).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae; Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88274327; PubMed=2839594;
RA McGeoch D.J., Dairymple M.A., Davison A.J., Dolan A., Frame M.C.,

```


RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
RA Dhami P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,
RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,
RA Gilbert J.G.R., Goward M.E., Grafham D.W., Griffiths M.N.D., Hall C.,
RA Hall R.E., Hall-Tamlyn G., Heathcote R.V., Ho S., Holmes S.,
RA Hunt S.E., Jones M.C., Kerhaw J., Kimberley A.M., King A.,
RA Laid G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,
RA Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,
RA McElroy J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,
RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey V., Rogers L., Ross M.T.,
RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
RA Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.,
RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,
RA Roe B.A., Chen F., Fu Y., Hu P., Hua A., Kenton S., Do A., Do T.,
RA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Do A., Do T.,
RA Lewis J., Lewis S., Lin S.-P., Loh P., Mala J.E., Nguyen T., Pan H.,
RA Phan S., Qian Y., Ray L., Ren Q., Shauli S., Sloan D., Song L.,
RA Wang S., Wang Y., Wang Z., White J., Wallingham D., Wu H., Yao Z.,
RA Zhan M., Zhang G., Chiesoe S., Murray J., Miller N., Minx P.,
RA Fulton R., Johnson D., Benis G., Bentley D., Graves T., Hawkins J.,
RA Cordes M., Du Z., Fulton L., Goela D., Layman D., Bradshaw H., Bourne S.,
RA Hinds K., Kemp K., Latreille P., Layman D., Bradshaw H., Bourne S.,
RA Scheet P., Walker C., Wattle P., Layman D., Bradshaw H., Bourne S.,
RA Korf I., Bedell J.A., Hillier L., Wardis E., Waterston R., Wilson R.,
RA Emanuel B.S., Shaikh T., Kurahashi H., Saitta S., Budarf M.L.,
RA Mcdermid H.E., Johnson A., Wong A.C.C., Morrow B.E., Edelman L.,
RA Kim U.-J., Shizuya H., Simon M.I., Dumanski J.P., Peyrard M., Kedra D.,
RA Seroushi E., Franssen I., Tapia I., Bruder C.E., O'Brien K.P.,
RA Wilkinson P., Bodenteich A., Hartman K., Hu X., Khan A.S., Lane L.,
RA Tilahun Y., Wright H.,
RA "The DNA sequence of human chromosome 22."
RA Nature 402:489-495 (1999).
RA [3]
RA SEQUENCE FROM N.A.
RA TISSUE=Muscle;
RA Strausberg R.;
RA Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RA -1- FUNCTION: May be involved in p53-induced cell cycle arrest in G2/M
RA phase by interfering with microtubule rearrangements that are
RA required to enter mitosis. Overexpression delays G2/M phase
RA progression.
RA -1- SUBCELLULAR LOCATION: Cytoplasmic. Associated with microtubules.
RA -1- DEVELOPMENTAL STAGE: Expressed in G2/M phase. Not detected in
RA quiescent cells.
RA -1- PIM: Phosphorylated in mitosis (By similarity).
RA -----
RA This SWISS-PROT entry is copyright. It is produced through a collaboration
RA between the Swiss Institute of Bioinformatics and the EMBL outstation -
RA the European Bioinformatics Institute. There are no restrictions on its
RA use by non-profit institutions as long as its content is in no way
RA modified and this statement is not removed. Usage by and for commercial
RA entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
RA or send an email to license@isb-sib.ch).
RA -----
RA EMBL; AF223408; AAF31459.1; -
RA EMBL; AL031588; CAB38415.1; -
RA EMBL; AL022325; CAB63079.1; -
RA EMBL; BC006325; AAH06325.1; -
RA KW Microtubules; Phosphorylation.
RA DOMAIN 22 27 POLY-SER.
RA CONFLICT 259 259 V -> I (IN REF. 1).
RA CONFLICT 506 506 R -> W (IN REF. 2; CAB38415).
RA SEQUENCE 720 AA; 76614 MW; ACD91CCCD008A89C CRC64;
RA -----
RA Query Match 8.7%; Score 90; DB 1; Length 720;
RA Best Local Similarity 22.7%; Pred. No. 0.91;
RA Matches 34; Conservative 20; Mismatches 36; Indels 60; Gaps 6;

QY 68 PCESAPKPTVLCVSCQCPARLLP-----NAIGRVKWMPPNGDFRCIPDRYRAQV----- 119
DB 341 PANSSRPLSNISKSGRMGMFAMLRPALPAGPVGASSW-----QAKRVDVSE 385
QY 120 ----QLLCPGGAAP-----RSRKVLVASC---KCRRLT 146
DB 386 LAAEQLTAPPSASPTQPTPEGGGWLNSCAWESSQLNKTIRRDSCLSNKTVM 445
QY 147 RPHNSEKLP-----GPTARPKQGRKRP 172
DB 446 TPTNQFKPKFSGESPDSTPKLSRAQRP 475
RESULT 10
LSHB TRIVU
ID LSHB TRIVU STANDARD; PRT; 141 AA.
AC O46482;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lutropin beta chain precursor (Luteinizing hormone beta subunit) (LSH-
DE beta) (LSH-B) (LH-B).
GN LHB
OS Trichosurus vulpecula (Brush-tailed possum).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Metatheria; Diprotodontia; Phalangeridae; Trichosurus.
OX NCBI TaxID=9337;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary;
RA MEDLINE=98345424; PubMed=9680384;
RA Harrison G.A., Deane B.M., Cooper D.W.;
RT "cDNA cloning of luteinizing hormone subunits from brushtail possum
RT and red kangaroo";
RL Mamm. Genome 9:638-642 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Lawrence S.B., McNatty K.P., Fidler A.E.;
RA Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
RA -1- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING
RA THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.
RA -1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
RA CHAIN WHICH CONFERES BIOLOGICAL SPECIFICITY TO THYROTROPIN,
RA LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
RA -1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
RA FAMILY.
RA -----
RA This SWISS-PROT entry is copyright. It is produced through a collaboration
RA between the Swiss Institute of Bioinformatics and the EMBL outstation -
RA the European Bioinformatics Institute. There are no restrictions on its
RA use by non-profit institutions as long as its content is in no way
RA modified and this statement is not removed. Usage by and for commercial
RA entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
RA or send an email to license@isb-sib.ch).
RA -----
RA EMBL; AF017448; AAC96019.1; -
RA EMBL; AF090388; AAC63526.1; -
RA HSSP; P01233; 1XUL.
RA InterPro; IPR000359; Cys_knot.
RA InterPro; IPR001545; Gly_hormoneB.
RA Pfam; PF00007; Cys_knot; 1.
RA SMART; SM00068; GHB; 1.
RA PROSITE; PS00261; GLYC_HORMONE_BETA_1; 1.
RA PROSITE; PS00689; GLYC_HORMONE_BETA_2; 1.
RA Hormone; Signal; Glycoprotein.
FT SIGNAL 1 22
FT CHAIN 23 141 LUTROPIN BETA CHAIN.
FT DISULFID 30 73 BY SIMILARITY.
FT DISULFID 44 93 BY SIMILARITY.
FT DISULFID 47 131 BY SIMILARITY.
FT DISULFID 55 109 BY SIMILARITY.
FT DISULFID 59 111 BY SIMILARITY.

THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.

-1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN, LUTROPIN, FOLLITROPIN AND GONADOTROPIN.

-1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN FAMILY.

PIR: PNO139; PNO139.
HSP; P01233; 1XUL.
InterPro: IPR000359; Cys_knot.
InterPro: IPR002400; GF_cysknot.
InterPro: IPR001545; Gly_hormone.
Pfam: PF00007; Cys_knot; 1.
PRINTS: PR00438; GFCYSKNOT.
SMART: SM00068; GHb; 1.
PROSITE: PS00261; GLYCO_HORMONE_BETA_1; 1.
PROSITE: PS00689; GLYCO_HORMONE_BETA_2; 1.
Hormone; Glycoprotein.

DISULFID 9 57 BY SIMILARITY.
FT FT FT FT FT BY SIMILARITY.
DISULFID 23 72 BY SIMILARITY.
FT FT FT FT FT BY SIMILARITY.
DISULFID 34 88 BY SIMILARITY.
FT FT FT FT FT BY SIMILARITY.
DISULFID 36 90 BY SIMILARITY.
FT FT FT FT FT BY SIMILARITY.
DISULFID 93 100 N-LINKED (GLNAC...)
FT CARBOHYD 13 13
FT SEQUENCE 118 AA; 12414 MW; 039F229BFC480F5D CRC64;

Query Match 8.8%; Score 91; DB 1; Length 118;
Best Local Similarity 29.3%; Pred. No. 0.12;
Matches 36; Conservative 12; Mismatches 57; Indels 18; Gaps

QY 55 CRELHYRFLTDGCRSAKPVTELVCSGQC-GPARLLPNAIGRVKWRPNGDPDFRCIPDR 113
Db 9 CRPINATLAZBZACPVCIITFTTSCAGYCPSMWRVLPAL-----PPVPZPVCTYRZ 61

QY 114 YRAQVRQLL-CPGGAAPR-SRKVRLVASCKKRLRFRHNOSELMDGFGFETARPKQ-GRKP 170
Db 62 LRFASIRLPGCPGVBWVFWFVALSCHGCPCELS-----SBCGFGRAZPLACBRSP 114

QY 171 RPG 173
Db 115 RPG 117

RESULT 9
GTSE HUMAN STANDARD; PRT; 720 AA.
AC ID GTSE HUMAN QSNV23; Q9BRE0; Q9V557; Q9UG29;
AD CTSY23; Q9BRE0; Q9V557; Q9UG29;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE G2 and S phase expressed protein 1 (B99 homolog).
DN GTSE1.
OS Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ON NCBI_Taxid=9606;
OX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=20057165; PubMed=10591208;
RX Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
RN [2]
RN SEQUENCE FROM N.A.
RP MEDLINE=20057165; PubMed=10591208;
RX Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
RN [2]
RN SEQUENCE FROM N.A.
RP MEDLINE=20057165; PubMed=10591208;
RX Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
RN [2]
RN SEQUENCE FROM N.A.
RP MEDLINE=20057165; PubMed=10591208;
RX Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
RN [2]
RN SEQUENCE FROM N.A.
RP MEDLINE=20057165; PubMed=10591208;
RX Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
RN [2]
RN SEQUENCE FROM N.A.
RP MEDLINE=20057165; PubMed=10591208;
RX Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
RN [2]
RN SEQUENCE FROM N.A.
RP MEDLINE=20057165; PubMed=10591208;
RX Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
RN [2]
RN SEQUENCE FROM N.A.
RP MEDLINE=20057165; PubMed=10591208;
RX Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
RN [2]
RN SEQUENCE FROM N.A.
RP MEDLINE=20057165; PubMed=10591208;
RX Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
RN [2]
RN SEQUENCE FROM N.A.
RP MEDLINE=20057165; PubMed=10591208;
RX Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
RN [2]
RN SEQUENCE FROM N.A.
RP MEDLINE=20057165; PubMed=10591208;
RX Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
RN [2]
RN SEQUENCE FROM N.A.
RP MEDLINE=20057165; PubMed=10591208;
RX Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
RN [2]
RN SEQUENCE FROM N.A.
RP MEDLINE=20057165; PubMed=10591208;
RX Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
RN [2]
RN SEQUENCE FROM N.A.
RP MEDLINE=20057165; PubMed=10591208;
RX Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
RN [2]
RN SEQUENCE FROM N.A.
RP MEDLINE=20057165; PubMed=10591208;
RX Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
RN [2]
RN SEQUENCE FROM N.A.
RP MEDLINE=20057165; PubMed=10591208;
RX Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
RN [2]
RN SEQUENCE FROM N.A.
RP MEDLINE=20057165; PubMed=10591208;
RX Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,

RA Beighton P., Mulligan J.T.,
 RT "Bone dysplasia sclerosteosis results from loss of the SOST gene
 RL product, a novel cysteine knot-containing protein.",
 CC Am. J. Hum. Genet. 68:577-589(2001).
 CC -!- FUNCTION: Seems to play a role in bone homeostasis (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
 CC -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF326738; AAK13453.1; -;
 DR InterPro: IPR000359; Cys_knot.
 DR SMART: SM00041; CT; 1.
 DR PROSITE: PS01185; CTCK_1; FALSE_NEG.
 DR PROSITE: PS01225; CTCK_2; FALSE_NEG.
 KW Glycoprotein.
 FT NON TER 1 1
 FT DOMAIN 51 141 CTCK.
 FT CARBOHYD 22 22 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT NON TER 176 176
 SQ SEQUENCE 176 AA; 19743 MW; 35F0CA61A25F4DB CRC64;

Query Match 81.3%; Score 839; DB 1; Length 176;
 Best Local Similarity 86.4%; Pred. No. 7e-69;
 Matches 152; Conservative 11; Mismatches 9; Indels 4; Gaps 2;

QY 8 NDATEVPLGLGEYEPPEPE-NNQTMRAENGRRPHHPYDAKDVSEVSCRELHYTRFLTD 66
 DB 1 NDATEIPLGEYEPPEPELNNKNTMRAENGRRPHHPYDAKDVSEVSCRELHYTRFLTD 60
 QY 67 GPCHSAKPVTVLVCSCGCGPARLLPNAIGRWKWRNPGDFRCIPDRYRAQRVQLLCPGG 126
 DB 61 GPCHSAKPVTVLVCSCGCGPARLLPNAIGRWKWRNPGDFRCIPDRYRAQRVQLLCPGG 120
 QY 127 AAPRSKRVLVASCKKRLTRFNQSLKDFGPETARPKGRKRP-CAKANQA 179
 DB 121 AAPRAKRVLVASCKKRLTRFNQSLKDFGPETARPKGRKRP-CAKANQA 176

RESULT 6
 LSHB_PHYCA
 ID LSHB_PHYCA STANDARD; PRT; 118 AA.
 AC P25330;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lutropin beta chain (luteinizing hormone beta subunit) (LSH-beta)
 DE (LSH-B) (LH-B).
 GN LHB.
 OS Physeter catodon (Sperm whale) (Physeter macrocephalus).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Cetacea; Odontoceti;
 OC Physeteridae; Physeter.
 OX NCBI_TaxID=9755;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=87032654; PubMed=3771098;
 RA Pankov Y.A., Karasyov V.S.;
 RT "Primary structure of sperm whale luteinizing hormone.";
 RL Int. J. Pept. Protein Res. 28:124-129(1986).
 RN [2]
 RP SEQUENCE.
 RX MEDLINE=84281133; PubMed=6466737;
 RA Pankov Y.A., Karasev V.S.;
 RT "Luteinizing hormone of the sperm whale. Amino acid sequences of

RT reduced and carboxymethylated beta-subunits.",
 RL Biochimia 49:1004-1018(1984).
 CC -!- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING
 CC THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.
 CC -!- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
 CC CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,
 CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
 CC -!- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
 CC FAMILY.
 DR PIR: P01411; P01411.
 DR HSP: P01233; 1X0L.
 DR InterPro: IPR000359; Cys_knot.
 DR InterPro: IPR002400; GF_Cysknot.
 DR InterPro: IPR001545; Gly_hormoneb.
 DR Pfam: PF00007; Cys_knot_1.
 DR PRINTS: PR00438; GFCYSKNOT.
 DR SMART: SM00068; GHB; 1.
 DR PROSITE: PS00261; GLYCO_HORMONE_BETA_1; 1.
 DR PROSITE: PS00689; GLYCO_HORMONE_BETA_2; 1.
 KW Hormone; Glycoprotein.
 FT DISULFID 9 57
 FT DISULFID 23 72 BY SIMILARITY.
 FT DISULFID 26 110 BY SIMILARITY.
 FT DISULFID 34 38 BY SIMILARITY.
 FT DISULFID 38 90 BY SIMILARITY.
 FT DISULFID 93 100 BY SIMILARITY.
 FT CARBOHYD 13 13 N-LINKED (GLCNAC. . .).
 SQ SEQUENCE 118 AA; 12412 MW; 81177A56382F15E7 CRC64;

Query Match 9.0%; Score 93; DB 1; Length 118;
 Best Local Similarity 30.1%; Pred. No. 0.079;
 Matches 37; Conservative 13; Mismatches 55; Indels 18; Gaps 6;

QY 55 CRELHYTRFLTDGCHSAKPVTVLVCSCGCG-SPARLLPNAIGRWKWRNPGDFRCIPDR 113
 DB 9 CRPINATLAAQNZACPFVITFTTISICAGYCPSVVRVLPAL-----DPVPFVCTYRQ 61
 QY 114 YRAQRVQLL-CPGAPAR-SRKVLVASCCKRLTRFNQSLKDFGPETARPK-GRKP 170
 DB 62 LRFASIRLPCGPPGVNPMVSPFVALSCHGCPCLSS-----SDCGPGRAPLACNRSP 114
 QY 171 RPG 173
 DB 115 RPG 117

RESULT 7
 LSHB_MELGA
 ID LSHB_MELGA STANDARD; PRT; 159 AA.
 AC P45646;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lutropin beta chain precursor (Luteinizing hormone beta subunit) (LSH-
 DE beta) (LSH-B) (LH-B).
 GN LHB.
 OS Meleagris gallopavo (Common turkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Meleagridae; Meleagris.
 OX NCBI_TaxID=9103;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pituitary.
 RX MEDLINE=95290073; PubMed=7772235;
 RA You S., Foster L.K., Silsby J.L., el Halawani M.E., Foster D.N.;
 RT "Sequence analysis of the turkey LH beta subunit and its regulation
 RT by gonadotropin-releasing hormone and prolactin in cultured
 RT pituitary cells.";
 RL J. Mol. Endocrinol. 14:117-129(1995).
 CC -!- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING
 CC THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.
 CC -!- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
 CC CHAIN WHICH CONFERS BIOLOGICAL SPECIFICITY TO THYROTROPIN,

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF326742; AAK13457.1; -
CC InterPro; IPR000359; Cys_knot.
CC SMART; SM00041; CT; 1.
CC PROSITE; PS01185; CTCK_1; FALSE_NEG.
CC PROSITE; PS01225; CTCK_2; FALSE_NEG.
CC Signal; Glycoprotein.
CC SIGNAL; 1 23 POTENTIAL.
CC CHAIN 24 213 SCLEROSTIN.
CC DOMAIN 82 172 CTCK.
CC CARBOHYD 53 53 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 175 175 N-LINKED (GLCNAC. .) (POTENTIAL).
CC SEQUENCE 213 AA; 23908 MW; 6DA7B5EDB674728A CRC64;
CC
CC Query Match 92.2%; Score 951.5; DB 1; Length 213;
CC Best Local Similarity 90.0%; Pred. No. 6.1e-79;
CC Matches 171; Conservative 10; Mismatches 4; Indels 5; Gaps 2;
CC
CC QY 1 QGWAQFNDATEVIGLGEYPPPPP--ENNQTMNRAENGRRPHHPHYDAKDYSEYSCREL 58
CC DB 24 QGWAQFNDATEIIPELGEYPPPPPELNNKTMRNRAENGRRPHHPFETKDYSEYSCREL 83
CC
CC QY 59 HYTRFLTDGPCRSAPKPVTELVCSCGQPARLLPNAIGRVKWRPNPDPFCIPDRYRAQR 118
CC DB 84 HFTRYVTDGPCRSAPKPVTELVCSCGQPARLLPNAIGRVKWRPNPDPFCIPDRYRAQR 143
CC
CC QY 119 VOLLCPGGAAPRSKRVLVASCKKRLTRFHNOSELKDFGPETARPKGRKPRP---GAK 175
CC DB 144 VOLLCPGGAAPRSKRVLVASCKKRLTRFHNOSELKDFGPETARPKGRKPRP---GAK 203
CC
CC QY 176 ANQAELENAY 185
CC DB 204 ANQAELENAY 213
CC
CC RESULT 4
CC SOST_HUMAN STANDARD; PRT; 213 AA.
CC AC Q9BOB4;
CC DT 15-JUN-2002 (Rel. 41, Created)
CC DT 15-JUN-2002 (Rel. 41, Last sequence update)
CC DE Sclerostin precursor.
CC GN SOST.
CC OS Homo sapiens (Human).
CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
CC OX NCBI_TaxID=9606;
CC [1]
CC SEQUENCE FROM N.A.
CC RX PubMed=11179006;
CC RA Balmain W., Ebelling M., Patel N., van Hul E., Olson P., Dioszegi M.,
CC Laezza C., Wuyts W., van den Ende J., Willems P., Paes-Alves A.F.,
CC Hill S., Bueno M., Ramos F.J., Tacconi P., Dikkers F.G., Stratakis C.,
CC Lindpaintner K., Vickery B., Foerzler D., van Hul W.;
CC "Increased bone density in sclerostosis is due to the deficiency of a
CC novel secreted protein (SOST).";
CC Hum. Mol. Genet. 10:537-543 (2001).
CC [2]
CC SEQUENCE FROM N.A.
CC RX PubMed=11179006;
CC RA Brunkow M.E., Gardner J.C., Van Ness J., Paepers B.W., Kovacevich B.R.,
CC Prohl S., Skonier J.E., Zhao L., Sabo P.J., Fu Y.H., Altsch R.S.,
CC Gillett L., Colbert T., Tacconi P., Galas D., Hamersma H.,
CC Beighton P., Mulligan J.T.;
CC "Bone dysplasia sclerostosis results from loss of the SOST gene
CC product, a novel cysteine knot-containing protein.";
CC Am. J. Hum. Genet. 68:579-589 (2001).
CC -!- FUNCTION: Seems to play a role in bone homeostasis.
CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
CC -!- TISSUE SPECIFICITY: Widely expressed at low levels with highest

```

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AF326740; AAK13455.1; -
 DR EMBL; AK017295; BAB30678.1; -
 DR EMBL; AF326737; AAK13452.1; -
 DR MGD; MGI:1921749; SOST.
 DR InterPro; IPR000359; Cys_knot.
 DR PROSITE; PS01185; CTCK_1; FALSE_NEG.
 DR PROSITE; PS01225; CTCK_2; FALSE_NEG.
 KW Signal; Glycoprotein.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 211 SCLEROSTIN.
 FT DOMAIN 80 170 CTCK.
 FT CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 173 173 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 72 72 G -> D (IN REF. 2).
 SQ SEQUENCE 211 AA; 23443 MW; AEB094E358E34961 CRC64;

Query Match 98.2%; Score 1013.5; DB 1; Length 211;
 Best Local Similarity 97.9%; Pred. No. 1.6e-84;
 Matches 184; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY 1 QGWAFRNDATVEIPGLGEYPEPPENNOTMNAENGRRPHHPYDAKDVSYSRELHY 60
 DB 24 QGWAFRNDATVEIPGLGEYPEPPENNOTMNAENGRRPHHPYDAKDVSYSRELHY 83
 QY 61 TRPLTDGPCRSAPVTTELVCSCGCGPARLLPNAIGRVKWRPNPGRFCIPDRYRAQRVQ 120
 DB 84 TRPLTDGPCRSAPVTTELVCSCGCGPARLLPNAIGRVKWRPNPGRFCIPDRYRAQRVQ 143
 QY 121 LCPGGAAPRSRKRLVASCCKKRLTRFNQSELKDFGPETARPQGRKPRP---GAKAN 177
 DB 144 LCPGGAAPRSRKRLVASCCKKRLTRFNQSELKDFGPETARPQGRKPRP---GAKAN 203
 QY 178 QAELENAY 185
 DB 204 QAELENAY 211

RESULT 2
 SOST_RAT
 ID SOST_RAT STANDARD; PRT; 213 AA.
 AC Q99567;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE Sclerostin precursor.
 GN SOST.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley;
 RX MEDLINE=21090529; PubMed=11179006;
 RA Brunkow M.E., Gardner J.C., Van Ness J., Paepers B.W., Kovacevich B.R.,
 RA Proll S., Skonier J.E., Zhao L., Sabo P.J., Fu Y.H., Alisch R.S.,
 RA Gillett L., Colbert T., Tacconi P., Galas D., Hamersma H.,
 RA Beighton P., Mulligan J.T.;
 RT "Bone dysplasia sclerosteosis results from loss of the SOST gene
 product, a novel cysteine knot-containing protein.";
 RL Am. J. Hum. Genet. 68:577-589(2001).
 CC -!- FUNCTION: Seems to play a role in bone homeostasis (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
 CC -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AF326741; AAK13456.1; -
 DR InterPro; IPR000359; Cys_knot.
 DR PROSITE; PS01185; CTCK_1; FALSE_NEG.
 DR PROSITE; PS01225; CTCK_2; FALSE_NEG.
 KW Signal; Glycoprotein.
 FT SIGNAL 1 28 POTENTIAL.
 FT CHAIN 29 213 SCLEROSTIN.
 FT DOMAIN 82 172 CTCK.
 FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 213 AA; 23974 MW; 6C56C878BCDC684B CRC64;

Query Match 95.2%; Score 982.5; DB 1; Length 213;
 Best Local Similarity 94.2%; Pred. No. 9.9e-82;
 Matches 179; Conservative 3; Mismatches 3; Indels 5; Gaps 2;

QY 1 QGWAFRNDATVEIPGLGEYPEPPPP--ENNQTMNAENGRRPHHPYDAKDVSYSREL 58
 DB 24 QGWAFRNDATVEIPGLREYPEPPQELLENNQTMNAENGRRPHHPYDTKDVSEYSEL 83
 QY 59 HYTFELTDGPCRSAPVTTELVCSCGCGPARLLPNAIGRVKWRPNPGRFCIPDRYRAQR 118
 DB 84 HYTFELTDGPCRSAPVTTELVCSCGCGPARLLPNAIGRVKWRPNPGRFCIPDRYRAQR 143
 QY 119 VOLLCPGGAAPRSRKRLVASCCKKRLTRFNQSELKDFGPETARPQGRKPRP---GAK 175
 DB 144 VOLLCPGGAAPRSRKRLVASCCKKRLTRFNQSELKDFGPETARPQGRKPRP---GAK 203
 QY 176 ANQAELENAY 185
 DB 204 ANQAELENAY 213

RESULT 3
 SOST_CERAE
 ID SOST_CERAE STANDARD; PRT; 213 AA.
 AC Q9BG78;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Sclerostin precursor.
 GN SOST.
 OS Cercopithecus aethiops (Green monkey) (Grivet).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Cercopithecus.
 OX NCBI_TaxID=9534;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21090529; PubMed=11179006;
 RA Brunkow M.E., Gardner J.C., Van Ness J., Paepers B.W., Kovacevich B.R.,
 RA Proll S., Skonier J.E., Zhao L., Sabo P.J., Fu Y.H., Alisch R.S.,
 RA Gillett L., Colbert T., Tacconi P., Galas D., Hamersma H.,
 RA Beighton P., Mulligan J.T.;
 RT "Bone dysplasia sclerosteosis results from loss of the SOST gene
 product, a novel cysteine knot-containing protein.";
 RL Am. J. Hum. Genet. 68:577-589(2001).
 CC -!- FUNCTION: Seems to play a role in bone homeostasis (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
 CC -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

A;Cross-references: EMBL:AF061947; NID:G3323606; PID:G3323607; PIDN:AAC40176.1

Query Match 7.9%; Score 82; DB 2; Length 2182;
Best Local Similarity 22.8%; Pred. No. 43;
Matches 50; Conservative 17; Mismatches 68; Indels 84; Gaps 9;

QY 4 QAFRDATEVIP-----CLGEYPEPP-----PNNQTNRAENG-----R 39
DB 1693 KASPEDQESLPHKPLADSGPGPEGKVGFLHQLPVATDTRDNTDQGGEPKDKER 1752
QY 40 PPHHPVDKDVSEYSCR-ELHYT-RFLTDGCRSAKP-----VTELVCSGCGPPAR 88
DB 1753 PPGVTETMDTGETAARHPDLEPTPLLPGRPPDRDGPESRBAELSLEELSISTRQPPAP 1812
QY 89 LLPNAIGRVKWRPNPGDFRCIPDRYRAQRVQLLCGGAAPRSRKVRLVASCCKRLTRF 148
DB 1813 LVPSPV-----TPTTAAPTMTGARAAGH-----1835
QY 149 HNOSELKDFGPETA--RPQGRKPRPGAKANOAELENAY 185
DB 1836 -----PEEAPPRPNRKRKLQDTEGKTLILDAY 1864

RESULT 11

S71793 head-inducing factor Cerberus - African clawed frog

C;Species: Xenopus laevis (African clawed frog)

C;Date: 12-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 21-Jul-2000

C;Accession: S71793

R;Boumeester, T.; Kim, S.H.; Sasaki, Y.; Lu, B.; de Robertis, E.M.

A;Title: Cerberus is a head-inducing secreted factor expressed in the anterior endoderm

A;Reference number: S71793; MUID:96338220; PMID:8757128

A;Accession: S71793

A;Status: preliminary; nucleic acid sequence not shown

A;Molecule type: mRNA

A;Residues: 1-270 <BOU>

A;Cross-references: EMBL:U64831; NID:G1513087; PIDN:AAC60012.1; PID:G1513088

Query Match

Best Local Similarity 7.8%; Score 80.5; DB 2; Length 270;

Matches 28; Conservative 27; Mismatches 53; Indels 25; Gaps 7;

QY 38 GRPPH--HPYDAKDVSRELYHTRFLTDGCRSAKPVTCLVCSGCGCPARLLPNAIG 95
DB 150 GAPQNTSHGSKAQEIEMKACKLTPFTQNLVHNC--DRMVIONNLCFGKC-----197

QY 96 RVKWRPNPGDFR-----CIPDRYRAQRVQLLCGGAAPRSRKVRLVASCCK-RLTRFH 149

DB 198 -ISLHVPNOODRRNTCSHCLPSKFTLNHLNCT-GSKNVVVKVMVVEECTCEAHKSNFH 255

QY 150 NQSELKDFGPETA 162

DB 256 QTAQ---FNWDTTS 265

RESULT 12

P4XRBV

minor inner core protein VP4 - bluetongue virus (serotype 10, American isolate)

C;Species: bluetongue virus

C;Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 10-Nov-1995

C;Accession: A26862; S10537

R;Yu, Y.; Fukusho, A.; Roy, P.

Nucleic Acids Res. 15, 7206, 1987

A;Title: Nucleotide sequence of the VP4 core protein gene (M4 RNA) of US bluetongue virus

A;Reference number: A26862; MUID:88015582; PMID:2921504

A;Accession: A26862

A;Molecule type: genomic RNA

A;Residues: 1-654 <YUY>

R;Roy, P.; Marshall, J.J.A.; French, T.J.

Curr. Top. Microbiol. Immunol. 162, 43-87, 1990

A;Title: Structure of the bluetongue virus genome and its encoded proteins.

A;Reference number: S10534; MUID:90345726; PMID:2166648

A;Accession: S10537

A;Status: preliminary

A;Molecule type: genomic RNA

A;Residues: 1-654 <ROY>

C;Genetics:

A;Map position: segment 4

C;Superfamily: bluetongue virus core protein VP4

C;Keywords: core protein

Query Match

Best Local Similarity 7.8%; Score 80; DB 1; Length 654;

Matches 53; Conservative 16; Mismatches 73; Indels 90; Gaps 9;

QY 8 NDATVPIGLGEYPPPPNNQTNRAENGGRPPHHPYDAKDV-----EYSRELHYTR 62
DB 141 NDANPELHKIYGMPIPP---LYMEYAEIGTRFDEPTDEKLVSMLHYIYVSAEYHYI- 196

QY 63 FLTDGCRSAKPVTCLVCSGCGCPARLL-----PNAIGRVKM--WRPNPGDFRCIPDR 113

DB 197 -----CGDLRLTLMQFKKSPGRFRVRLWHYVDIAPB--CSDPN 234

QY 114 YRAQRVQLLCGGAAPRSRKVRLVASCCKRLTR-----FHNQSELKDFGPETARPQKG 167

DB 235 VIVHNIW-----DSKDKILKMNFLKRVPRPFINDVSSDRSQNDHEWETTRFAD 286

QY 168 RK-----PRPGAKANOAELEN 183

DB 287 RLGEIAYEMGAFSSALIKHRIPNSKDEYHCISTYLFQPGADADNYELRN 338

RESULT 13

AS7534 mucin SAC (clone L31) - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change 20-Apr-2000

C;Accession: A57534

R;Lesuffleur, T.; Roche, F.; Hill, A.S.; Lacasa, M.; Fox, M.; Swallow, D.M.; Zweibaum, A.

J. Biol. Chem. 270, 13665-13673, 1995

A;Title: Characterization of a mucin cDNA clone isolated from HT-29 mucus-secreting cells

A;Reference number: A57534; MUID:95293957; PMID:7775418

A;Accession: A57534

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-1042 <LES>

A;Cross-references: GB:Z46314; NID:G1052607; PIDN:CAA68307.1; PID:G1052608

C;Genetics:

A;Gene: GDB:MUC5AC

A;Cross-references: GDB:454136; OMIM:158373

A;Map position: l1p15.5-11p15.5

C;Superfamily: von Willebrand factor type C repeat homology

F;678-746/Domain: von Willebrand factor type C repeat homology <WVC>

Query Match

Best Local Similarity 7.8%; Score 80; DB 2; Length 1042;

Matches 35; Conservative 16; Mismatches 55; Indels 58; Gaps 6;

QY 23 PPEPNNQTNRAENGGRPPHHPYDAKDVSEYSCRELHYHTRFLTDGCRSAKPVTCLVCSG 82

DB 819 PYPYQNS-----TCAYTHRSLLIQOQCSSESSEPVRLAYCRG 855

QY 83 QCGPARLLPNAIGRVKWRPNPGDFR---CIPDRYRAQRVQLLCGGAAPRSRKVRLVAS 139

DB 856 NCGDSSSMYSLEG-----NTVEHRCQCCQLRTSLRNVTLCITDSS-RAFSYTEVEE 907

QY 140 CKCKRLTRFHNQSELKDFGPETARPQGRK-PRPGAKANOAELE 182

DB 908 CGC-----MGRCPAGDPDQHQSEEA 928

RESULT 14

I51242

luteinizing hormone beta-subunit - quail

C;Species: Coturnix coturnix (quail)

Mon Mar 31 09:44:36 2003

us-09-867-274-4.xpr

```

A;Accession: T04572
A;Molecule type: DNA
A;Residues: 1-334 <BEV>
A;Cross-references: EMBL:AL021635
A;Experimental source: cultivar Columbia; BAC clone T12H17
C;Genetics:
A;Map position: 4
A;Introns: 139/1; 161/3; 205/3
A;Note: T12H17.160

Query Match      8.0%; Score 82.5; DB 2; Length 334;
Best Local Similarity 25.9%; Pred. No. 6;
Matches 28; Conservative 13; Mismatches 30; Indels 37; Gaps 4;

QY 21 PEPPENNQTMNRAENG-----GRPPHPYDADKDVSEYSCRELHYTRFLTDGPCR 70
Db 49 PPPPQNSFTPSAAMDGFSSGPIKRRGRPRKYGHGAAVT-----LSPNPIS 96

QY 71 SAKPVTELVC-----GCGPARLLPNAIGRVKMW-----WRPN 103
Db 97 SAAPTSHVIDFTSTSEKRGKMKPATPTSSFIRPKYQVENLGWSPS 144

RESULT 9
WMBEF9
UL49 protein - human herpesvirus 1 (strain 17)
C;Species: human herpesvirus 1
C;Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 16-Jun-2000
C;Accession: D30089
R;McGeoch, D.J.; Dalrymple, M.A.; Davison, A.J.; Dolan, A.; Frame, M.C.; McNab, D.; Perrett, J.; Gen. Virol. 69, 1531-1574, 1988
A;Title: The complete DNA sequence of the long unique region in the genome of herpes simi
A;Reference number: A30083; MUID:88274327; PMID:2839594
A;Accession: D30089
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-301 <MCG>
A;Cross-references: GB:X14112; NID:G1944536; PIDN:CAA32299.1; PID:G59549; GB:D00317
C;Genetics:
A;Gene: UL49
C;Superfamily: varicella-zoster virus gene 9 protein

Query Match      7.9%; Score 82; DB 1; Length 301;
Best Local Similarity 26.1%; Pred. No. 6;
Matches 43; Conservative 17; Mismatches 67; Indels 38; Gaps 10;

QY 21 PEPPENNQ---TMNRAENGRRPPHPYDADKDVSEY-----SCRELHYTRFLTDGPCRSAK 73
Db 33 PDSPPDTSRRGALQTKSRORGEVRFVQYDESDYALYGSSSEDEH-----PEVP-RTRR 86

QY 74 PVTELVCSGCGPARLLP-----NAIGRVKWRPNPGDFRCIPDRYRAQVQLLCPGGAA 128
Db 87 PVSGAVLSGP-GPARAPPPAGSGGAGRTPTTAPRAP-----RTQRVATKAP--AA 134

QY 129 PRSRKRLVASCKKCLTRFNHNSKELKDFGPETARPKGKPRPG 173
Db 135 PAETTR-----GRKSAQPESALPD-APASTATPRSKTTPAQG 171

RESULT 10
T14320
calcineurin inhibitor cain - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T14320
J;Lai, M.M.; Burnett, P.E.; Wolosker, H.; Blackshaw, S.; Snyder, S.H.
J. Biol. Chem. 273, 18325-18331, 1998
A;Title: Cain, a novel physiologic protein inhibitor of calcineurin.
A;Reference number: Z17979; MUID:98325042; PMID:9660798
A;Accession: T14320
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-2182 <LAI>

A;Introns: 294/2
A;Superfamily: Escherichia coli ribosomal protein L2
C;Keywords: mitochondrion; protein biosynthesis; ribosome

Query Match      8.5%; Score 88; DB 2; Length 332;
Best Local Similarity 25.5%; Pred. No. 1.9;
Matches 27; Conservative 16; Mismatches 35; Indels 28; Gaps 5;

QY 3 WOAFRNDATEVIGLGYPPPPPPNNQTMNRAENGRRPPHPYDADKDVSEYSCRELHYTR 62
Db 244 WLSFRROFTDLGVAAHNSKPKTDQ-----AKPIGEGT----- 282

QY 63 FLTGDPCR-SAKPVTELVCSCGCGPARLLPNAIGRVKWRPNPGPDF 107
Db 283 --KOGACKVDRAPTYIIASHOLEAGRWVNC-----DWSKPSSTDF 322

RESULT 7
S29717
adenylate cyclase (EC 4.6.1.1) type 5 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 25-Feb-1994 #sequence_revision 01-Sep-1995 #text_change 04-Sep-1998
C;Accession: S29717; B47202
R;Glatt, C.E.; Snyder, S.H.
Nature 361, 536-538, 1993
A;Title: Cloning and expression of an adenylyl cyclase localized to the corpus striatum.
A;Reference number: S29717; MUID:93156838; PMID:8429907
A;Accession: S29717
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1223 <GLA>
R;Premont, R.T.; Chen, J.; Ma, H.W.; Ponnappalli, M.; Iyengar, R.
Proc. Natl. Acad. Sci. U.S.A. 89, 9809-9813, 1992
A;Title: Two members of a widely expressed subfamily of hormone-stimulated adenylyl cycl
A;Reference number: A47202; MUID:93028552; PMID:1409703
A;Accession: B47202
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 91,'GKG','95-97,'E','134-183,'R','185-209,'V','211-237,'A','239-346,'I','348-511,'
RE>
A;Experimental source: liver, kidney
A;Note: sequence extracted from NCBI backbone (NCBIP:115852)
C;Superfamily: human adenylyl cyclase; guanylate cyclase catalytic domain homology
C;Keywords: phosphorus-oxygen lyase
F;373-608/Domain: guanylate cyclase catalytic domain homology <GCC>
F;981-1220/Domain: guanylate cyclase catalytic domain homology <GCC2>

Query Match      8.2%; Score 84.5; DB 2; Length 1223;
Best Local Similarity 27.9%; Pred. No. 14;
Matches 41; Conservative 11; Mismatches 54; Indels 41; Gaps 7;

QY 10 ATEVIGP-----IGVEPEPPENNQTMNRAENGRRPPHPYDADKDVSEYSCRELHY 60
Db 2 ATPPTPGDRPAAPPSDLGE-PVTPQOORLASRSGGDDDEDDPLSGDDPLEASASAVL 60

QY 61 TRFLTGDGCRSAKPVTELVCSCGCGPARLLP-----AIGRVKWRPNPGPDF-- 107
Db 61 SR-----FGRSAATTAAGGSGGARLAGGSGTRAPLRAAVAVRR-----RPQOPQVAR 111

QY 108 RCIPDRYR-----AQRVQLCPG 125
Db 112 RCAPARWRWNRVSEKAERPRELEPG 138

RESULT 8
T04572
hypothetical protein T12H17.160 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 14-May-1999
A;Accession: T04572
R;Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duisterhoef, A.; Bancroft,
submitted to the Protein Sequence Database, February 1998
A;Reference number: Z15377

```

A;Genome: mitochondrion

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:13:46 ; Search time 26.64 Seconds
(without alignments)
667.600 Million cell updates/sec

Title: US-09-867-274-4

Perfect score: 1032

Sequence: 1 CQWQAFRNDATVIFGLGEY.....KGRKPRPGAKANQALENAY 185

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73.*

1: piri.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	249	24.1	134	2 T08710	hypothetical prote
2	93	9.0	118	2 PNO141	lutropin beta chai
3	92.5	9.0	159	2 I51373	lutropinizing horm
4	91	8.8	118	2 PNO139	lutropin beta chai
5	89.5	8.7	158	2 A61091	lutropin beta chai
6	88	8.5	332	2 S46947	ribosomal protei
7	84.5	8.2	1223	2 S29717	adenylate cyclase
8	82.5	8.0	334	2 T04572	hypothetical prote
9	82	7.9	301	1 WMBEP9	UL49 protein - hum
10	82	7.9	2182	2 T14320	calcineurin inhibi
11	80.5	7.8	270	2 S71793	head-inducing fact
12	80	7.8	654	1 P4XRBV	minor inner core p
13	80	7.8	1042	2 A57534	mucin SAC (clone L
14	79.5	7.7	166	2 I51242	lutropinizing horm
15	79.5	7.7	601	2 T22025	hypothetical prote
16	79.5	7.7	601	2 D89711	protein P40E10.4 l
17	78.5	7.6	1142	2 T30272	hypothetical prote
18	78.5	7.6	1098	2 T18397	protein CTRP - mal
19	77.5	7.5	188	2 JC4680	vascular endotheli
20	77.5	7.5	349	2 S49606	ribosomal protei
21	77	7.5	834	2 T42702	hypothetical prote
22	76.5	7.4	165	1 KTHUB	choriognadotropin
23	76	7.4	232	2 A60083	neural induction h
24	76	7.4	405	2 S33601	cell adhesion mole
25	76	7.4	888	2 I58378	tyrosine kinase -
26	76	7.4	1469	2 B36665	slit protein kinase
27	75.5	7.3	3570	2 T45025	mucin MUC5B, trach
28	75	7.3	322	2 AH3011	thiorodoxin trxa l
29	75	7.3	331	2 G98272	probable thioredox

30 75 7.3 1736 2 T00391
31 74.5 7.2 128 2 S74085
32 74.5 7.2 275 2 T50578
33 74 7.2 178 2 I51824
34 74 7.2 178 2 A47291
35 74 7.2 269 2 D75631
36 74 7.2 294 2 T34537
37 74 7.2 379 2 A35669
38 74 7.2 866 2 T29197
39 73.5 7.1 145 2 I37231
40 73.5 7.1 402 2 T48729
41 73.5 7.1 623 2 B83399
42 73.5 7.1 650 2 A34498
43 73.5 7.1 701 2 T52384
44 73.5 7.1 825 2 JC4163
45 73 7.1 433 2 D84335

ALIGNMENTS

RESULT 1

T08710

hypothetical protein DKFZp564D206.1 - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 13-Aug-1999

C:Accession: T08710

R:Wambutt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, March 1999

A:Reference number: 216471

A:Accession: T08710

A:Molecule type: mRNA

A:Residues: 1-134 <WAM>

A:Cross-references: EMBL:AL050024

A:Experimental source: fetal brain; clone DKFZp564D206

C:Genetics:

A:Note: DKFZp564D206.1

Query Match 24.1%; Score 249; DB 2; Length 134;

Best Local Similarity 40.3%; Pred. No. 3.3e-15;

Matches 52; Conservative 28; Mismatches 43; Indels 6; Gaps 4;

QY 55 CSELHYRFLTDGCRSAKPVTELVCSCGCGPARLLPNAIG---RVKWW-RPNGPDFRCI 110
DB 3 CRELRSTKIYSDGQCTSIISPLKELVCAGECLLLPVLPNWIGGYGTYKWSRSSSEWECV 62
QY 111 PDRYAQRVQLCPGGAAPRRKRVLRVASCCKKRLTRFHNSQELKDFGFPETARP-QKGRK 169
DB 63 NDKTRQRIQLCQDQGST-RTYKTIIVTATACKRYTRQHNESHNFESMSPAKPVQHRE 121

QY 170 PRPGAKANQ 178
DB 122 RKRASKSSK 130

RESULT 2

PNO141

lutropin beta chain - sperm whale

N:Alternate names: luteinizing hormone beta chain

C:Species: Physeter catodon (sperm whale)

C>Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 07-May-1999

C:Accession: PNO141

R:Pankov, Y.A.; Karasev, V.S.

Biokhimiia 49, 1004-1018, 1984

A>Title: Luteinizing hormone of the sperm-whale: amino acid sequence of reduced and carb

A:Reference number: PNO141; MUID:84281133; PMID:6466737

A:Accession: PNO141

A:Molecule type: protein

A:Residues: 1-118 <PAM>

A:Note: article in Russian with English abstract

C:Superfamily: pituitary glycoprotein hormone beta chain

C:Keywords: glycoprotein; hormone

F;9-34,23-57,26-88,38-110,72-100,90-93/Disulfide bonds: #status predicted

Search completed: March 28, 2003, 14:16:14
Job time : 52.8 secs

XX WPI; 2000-412321/35.
 DR N-PSDB; AAA29056.
 XX
 XX Nucleic acids (I) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures
 XX
 XX Claim 3; Page 119-120; 162pp; English.
 XX
 XX This shows a variant human transforming growth factor-beta (TGF-beta)
 CC binding protein designated BEER V101, which comprises a substitution of
 CC isoleucine for the wild-type valine at residue 10. The cDNA and protein
 CC may be used for prevention, treatment and diagnosis of diseases
 CC associated with inappropriate BEER expression. For example, they may be
 CC used to treat disorders associated with decreased TGF-beta BP expression.
 CC The cDNA or vectors may be administered to treat diseases by rectifying
 CC mutations or deletions in a patient's genome that affect the activity of
 CC BEER by expressing inactive proteins or to supplement the patients own
 CC production of BEER polypeptides. The nucleic acids may be used for
 CC recombinant production of BEER, gene therapy, antisense therapy, as
 CC probes for diagnostic assays and for functional studies. BEER may be used
 CC to raise antibodies and for identification of BEER modulators. BEER
 CC antagonists may be used to increase bone mineral content for the
 CC treatment of disorders such as osteopenia, osteoporosis, fractures and
 CC other disorders associated with low mineral content.
 XX
 XX Sequence 213 AA;
 SQ
 Query Match 89.5%; Score 923.5; DB 21; Length 213;
 Best Local Similarity 87.9%; Pred. No. 2.9e-78;
 Matches 167; Conservative 10; Mismatches 8; Indels 5; Gaps 2;
 QY 1 QGWAQRNDATVTPGLGEYEPPEPP--ENNQTMNRAENGRRPHHPYDAKDVSEYSCREL 58
 DB 24 QGWAQFNKNDATVTPGLGEYEPPEPP--ENNQTMNRAENGRRPHHPYDAKDVSEYSCREL 83
 QY 59 HVTRELTDPGCRSAKPVTELVCSCGCGPARLLPNAIGRKWKWRPNPGDFRCIPDRYRAQR 118
 DB 84 HFTRYVTDGPCRSAPVTELVCSCGCGPARLLPNAIGRKWKWRPNPGDFRCIPDRYRAQR 143
 QY 119 VOLLCPGGAAPRSRKRVLVASCCKRLTRFHNSQLKDFGTEARPKQGRKPRP---GAK 175
 DB 144 VOLLCPGGAAPRSRKRVLVASCCKRLTRFHNSQLKDFGTEARPKQGRKPRP---GAK 203
 QY 176 ANQAELENAY 185
 DB 204 ANQAELENAY 213
 RESULT 14
 AAY96436
 ID AAY96436 standard; Protein; 213 AA.
 XX
 XX AAY96436;
 XX
 XX 12-SEP-2000 (first entry)
 XX
 XX Human TGF-beta binding protein (BEER) variant P38R.
 XX
 XX osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 XX BEER; variant; P38R; gene therapy; antisense therapy; fracture;
 XX chromosome 17q12-21; bone mineralization.
 XX
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH Misc-difference 38
 FT /label= P38R
 FT /note= "wild type proline has been substituted with
 FT arginine"
 XX
 XX W0200032773-A1.
 PN

XX 08-JUN-2000.
 XX 24-NOV-1999; 99WO-US27990.
 XX
 XX 27-NOV-1998; 98US-0110283.
 XX
 XX (DARW-) DARWIN DISCOVERY LTD.
 XX
 XX Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepers BW;
 PI Van Ness J, Winkler DG;
 XX
 XX WPI; 2000-412321/35.
 DR N-PSDB; AAA29062.
 XX
 XX Nucleic acids (I) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures
 XX
 XX Disclosure; Page 121; 162pp; English.
 XX
 XX This shows a variant human transforming growth factor-beta
 CC (TGF-beta) binding protein designated BEER P38R. The encoded protein
 CC comprises a substitution of arginine for the wild-type proline at
 CC residue 38. The cDNA and protein may be used for prevention, treatment
 CC and diagnosis of diseases associated with inappropriate BEER expression.
 CC For example, they may be used to treat disorders associated with
 CC decreased TGF-beta BP expression. The cDNA or vectors may be administered
 CC to treat diseases by rectifying mutations or deletions in a patient's
 CC genome that affect the activity of BEER by expressing inactive proteins
 CC or to supplement the patients own production of BEER polypeptides. The
 CC functional studies. BEER may be used to raise antibodies and for
 CC therapy, antisense therapy, as probes for diagnostic assays and for
 CC identification of BEER modulators. BEER antagonists may be used to
 CC increase bone mineral content for the treatment of disorders such as
 CC osteopenia, osteoporosis, fractures and other disorders associated with
 CC low mineral content.
 XX
 XX Sequence 213 AA;
 SQ
 Query Match 89.5%; Score 923.5; DB 21; Length 213;
 Best Local Similarity 87.9%; Pred. No. 2.9e-78;
 Matches 167; Conservative 10; Mismatches 8; Indels 5; Gaps 2;
 QY 1 QGWAQRNDATVTPGLGEYEPPEPP--ENNQTMNRAENGRRPHHPYDAKDVSEYSCREL 58
 DB 24 QGWAQFNKNDATVTPGLGEYEPPEPP--ENNQTMNRAENGRRPHHPYDAKDVSEYSCREL 83
 QY 59 HVTRELTDPGCRSAKPVTELVCSCGCGPARLLPNAIGRKWKWRPNPGDFRCIPDRYRAQR 118
 DB 84 HFTRYVTDGPCRSAPVTELVCSCGCGPARLLPNAIGRKWKWRPNPGDFRCIPDRYRAQR 143
 QY 119 VOLLCPGGAAPRSRKRVLVASCCKRLTRFHNSQLKDFGTEARPKQGRKPRP---GAK 175
 DB 144 VOLLCPGGAAPRSRKRVLVASCCKRLTRFHNSQLKDFGTEARPKQGRKPRP---GAK 203
 QY 176 ANQAELENAY 185
 DB 204 ANQAELENAY 213
 RESULT 15
 AAY96434
 ID AAY96434 standard; Protein; 176 AA.
 XX
 XX AAY96434;
 XX
 XX 12-SEP-2000 (first entry)
 XX
 XX Bovine TGF-beta binding protein (BEER).
 XX osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 XX

Query Match 90.4%; Score 932.5; DB 23; Length 213;
 Best Local Similarity 88.4%; Pred. No. 4.2e-79;
 Matches 168; Conservative 10; Mismatches 7; Indels 5; Gaps 2;

QY 1 QGWAQFRNDATVPGIGYVPEPPP--ENNQTMNRAENGRRPHHPYDAKDVSEYSCREL 58
 DB 24 QGWAQFKNDAATEIIPELGEYVPEPPPPELENNKTMNRAENGRRPHHPFETKDVSEYSCREL 83
 QY 59 HYTRFLTDGPCRSAPVTELVCSGQCGPARLLPNAIGRVKWRPNPGDFRCIPDRYRAQR 118
 DB 84 HTRYVTDGPCRSAPVTELVCSGQCGPARLLPNAIGRVKWRPNPGDFRCIPDRYRAQR 143
 QY 119 VOLLCPGGAAPRSRKVRLVASCKKRLTRFNQSELKDFGPETAPQKGRKPRP---GAK 175
 DB 144 VOLLCPGGEAPRARKVRLVASCKKRLTRFNQSELKDFGTEAARPKQGRKPRPARSAK 203
 QY 176 ANQAELENAY 185
 DB 204 ANQAELENAY 213

RESULT 12
 AAEL17089
 ID AAEL17089 standard; Protein; 213 AA.
 AC AAEL17089;
 DT 18-APR-2002 (first entry)
 DE Human osteolevin protein.
 KW Human; osteolevin; osteopathic; cytostatic; bone formation; osteoporosis;
 KW Van Buchem-sclerosteosis disease; sclerosteosis; transgenic animal;
 KW Paget's disease; chromosome 17.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Peptide 1..19
 FT /label= Signal_peptide
 FT Protein 20..213
 FT /label= Mature_osteolevin_protein
 FT Misc-difference 10
 FT /note= "During polymorphism wild type Val is
 substituted with Ile"
 PN WO200198491-A2.
 XX 27-DEC-2001.
 XX 15-JUN-2001; 2001WO-EP06795.
 XX 19-JUN-2000; 2000EP-0112867.
 XX (HOFF) HOFFMANN LA ROCHE & CO AG F.
 XX (UYIN-) UNIV INSTELLING ANTWERPEN UIA.
 XX Balenans W, Ebeling M, Foernzler D, Patel N, Van Hul W;
 XX Vickery BH;
 XX WPI; 2002-139789/18.
 XX N-PSDB; AAD27576, AAD27577.
 XX Novel genetic polymorphisms in the Van Buchem-sclerosteosis disease
 XX region that are associated with abnormal bone formation useful for
 XX diagnosis and assessment of osteoporosis or sclerosteosis in humans
 XX Claim 11; Fig 4; 70pp; English.
 XX The invention relates to a nucleic acid encoding osteolevin region
 XX polymorphisms. The invention also relates to genetic polymorphisms in
 XX the Van Buchem-sclerosteosis disease region that are associated with
 XX disorders resulting in either net excess bone formation or insufficient

CC bone formation in humans. Osteolevin DNA is useful for screening for
 CC osteolevin polymorphisms associated with abnormal bone formation in a
 CC subject and for the presence of a heritably linked form of abnormal bone
 CC formation in a subject, by determining the presence of a polymorphism in
 CC the osteolevin nucleic acid sequence obtained from the subject.
 CC Osteolevin protein is useful for treating diseases associated with
 CC abnormal bone formation, such as sclerosteosis, Van Buchem's disease and
 CC Paget's disease. Nucleic acids which encode genes in the osteolevin
 CC region or their modified forms can also be used to generate either
 CC transgenic animals or knockout animals which are useful in the screening
 CC and development of therapeutically useful reagents. Osteolevin proteins
 CC are useful in pharmacological characterisation of novel modulators of the
 CC activity of protein and protein complexes. Human osteolevin gene is
 CC located on chromosome 17. The present sequence is human osteolevin
 CC protein.
 XX
 SQ Sequence 213 AA;

Query Match 90.4%; Score 932.5; DB 23; Length 213;
 Best Local Similarity 88.4%; Pred. No. 4.2e-79;
 Matches 168; Conservative 10; Mismatches 7; Indels 5; Gaps 2;

QY 1 QGWAQFRNDATVPGIGYVPEPPP--ENNQTMNRAENGRRPHHPYDAKDVSEYSCREL 58
 DB 24 QGWAQFKNDAATEIIPELGEYVPEPPPPELENNKTMNRAENGRRPHHPFETKDVSEYSCREL 83
 QY 59 HYTRFLTDGPCRSAPVTELVCSGQCGPARLLPNAIGRVKWRPNPGDFRCIPDRYRAQR 118
 DB 84 HTRYVTDGPCRSAPVTELVCSGQCGPARLLPNAIGRVKWRPNPGDFRCIPDRYRAQR 143
 QY 119 VOLLCPGGAAPRSRKVRLVASCKKRLTRFNQSELKDFGPETAPQKGRKPRP---GAK 175
 DB 144 VOLLCPGGEAPRARKVRLVASCKKRLTRFNQSELKDFGTEAARPKQGRKPRPARSAK 203
 QY 176 ANQAELENAY 185
 DB 204 ANQAELENAY 213

RESULT 13
 AAY96430
 ID AAY96430 standard; Protein; 213 AA.
 AC AAY96430;
 DT 12-SEP-2000 (first entry)
 DE Human TGF-beta binding protein (BEER) variant V101.
 DE osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 DE BEER; variant; V101; gene therapy; antisenese therapy; fracture;
 DE chromosome 17q12-21; bone mineralization.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Misc-difference 10
 FT /label= V101
 FT /note= "wild type valine has been substituted with
 isoleucine"

WO200032773-A1.
 PD 08-JUN-2000.
 XX 24-NOV-1999; 99WO-US27990.
 XX 27-NOV-1998; 98US-0110283.
 XX (DARW-) DARWIN DISCOVERY LTD.
 XX Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepers BW;
 XX Van Ness J, Winkler IG;

16-JAN-2001; 2001US-261910P.
 16-JAN-2001; 2001US-261939P.
 16-JAN-2001; 2001US-262150P.
 16-JAN-2001; 2001US-264395P.
 02-FEB-2001; 2001US-266421P.
 02-FEB-2001; 2001US-267623P.
 08-FEB-2001; 2001US-267623P.
 08-FEB-2001; 2001US-267623P.
 03-MAR-2001; 2001US-274399P.
 03-MAR-2001; 2001US-280982P.
 04-APR-2001; 2001US-282129P.
 04-APR-2001; 2001US-282199P.
 09-MAY-2001; 2001US-290589P.
 23-MAY-2001; 2001US-2917092.
 01-JUN-2001; 2001US-2917800.
 20-JUN-2001; 2001US-2919692.
 29-JUN-2001; 2001US-2921066.
 09-JUL-2001; 2001US-2921735.
 XX (GETH) GENENTECH INC.
 XX Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC;
 PI Gurney AL, Smith V, Stephan J, Watanabe CK, Wood WI, Zhang Z;
 PI Fong S;
 XX WPI; 2002-362426/39.
 DR N-PSDB; ABK69992.
 XX New PRO polypeptides and polynucleotides encoding the polypeptides,
 PT useful in gene therapy, chromosome identification, tissue typing, or
 PT for genetic analysis of individuals with genetic disorders -
 XX Claim 11; Figure 64; 218pp; English.
 XX This invention relates to the cDNA and protein sequences of novel
 CC secreted and transmembrane polypeptides PRO polypeptides. The
 CC invention also comprises a method for producing the proteins of the
 CC invention by recombinant means and antibodies specific for the protein
 CC of the invention. The antibody may be used for detecting the PRO
 CC proteins of the invention and may be used to modify their activity.
 CC polynucleotides may be used as hybridisation probes for a cDNA library
 CC to isolate the full-length PRO cDNA or to isolate other cDNAs, to
 CC construct hybridisation probes for mapping the gene which encodes that
 CC PRO and for genetic analysis of individuals with genetic disorders, in
 CC assays to identify other proteins or molecules involved in binding
 CC reaction, to generate transgenic animals or knock-out animals which in
 CC turn are useful in the development and screening of therapeutically
 CC useful reagents, for chromosome identification, and tissue typing. The
 CC PRO polypeptides are useful in gene therapy, and as molecular weight
 CC markers for protein electrophoresis purposes. The sequences may
 CC also be used to detect overexpression on PRO polypeptides in cancerous
 CC tumours and for screening for differentially expressed genes using
 CC microarray technology. The present sequence represents a human PRO
 CC protein of the invention.
 SQ Sequence 213 AA;
 Query Match 90.4%; Score 932.5; DB 23; Length 213;
 Best Local Similarity 88.4%; Pred. No. 4.2e-79;
 Matches 168; Conservative 10; Mismatches 7; Indels 5; Gaps 2;
 QY 1 QGWAQFRNDATVEIPGLGYPPPPP--ENNQTMNRAENGRRPHHPYDKDVSYSCREL 58
 24 QGWAQFRNDATVEIPGLGYPPPPP--ENNQTMNRAENGRRPHHPYDKDVSYSCREL 83
 QY 59 HYTRFTDGCRAKVTYELVCSGGCPARLLPNAIGKWRPNPDPFCIPDRIYRAQR 118
 84 HFTRYVTDGFCRAKVTYELVCSGGCPARLLPNAIGKWRPNPDPFCIPDRIYRAQR 143
 QY 119 VQLLCGGAAPRKRKRLVASCCKKELTRFHNSQSELKDFGPTARPKQGRKRP--GAK 175
 144 VQLLCGGAAPRKRKRLVASCCKKELTRFHNSQSELKDFGPTARPKQGRKRP--GAK 203
 QY 176 ANQAELENAY 185

Db 204 ANQAELENAY 213
 RESULT 11
 ABB07209
 ID ABB07209 standard; Protein; 213 AA.
 XX ABB07209;
 AC ABB07209;
 XX 26-MAR-2002 (first entry)
 XX Human cloaked-2 polypeptide sequence.
 XX Cloaked-2; cysteine knot motif; nephrotropic; cardiant; immunomodulator;
 KW hepatotropic; antiinflammatory; antithyroid; cytostatic; neuroprotective;
 KW antianemic; hypotensive; antiarrhythmic; antiarteriosclerotic; muscular;
 KW antidiabetic; anorectic; gene therapy; cell therapy; antisense therapy;
 KW human.
 XX Homo sapiens.
 OS Key Location/Qualifiers
 FH Peptide 1..23
 FT /note= "signal peptide"
 FT Protein 24..213
 FT /note= "mature protein (ABB07207)"
 XX WC200192308-A2.
 XX 06-DEC-2001.
 XX 29-MAY-2001; 2001WO-US17478.
 XX 01-JUN-2000; 2000US-208550P.
 XX 04-AUG-2000; 2000US-223542P.
 XX (AMGE-) AMGEN INC.
 XX Paszty CU, Gao Y;
 XX WPI; 2002-114325/15.
 XX N-PSDB; ABA94293.
 XX New human and mouse cysteine-knot polypeptide designated as Cloaked-2,
 PT for treating or preventing kidney, heart (e.g. myocardial infarction)
 PT or liver (e.g. hepatitis) diseases -
 XX Example 1; Fig 1; 170pp; English.
 XX The invention relates to polypeptides comprising a cysteine knot motif
 CC and designated as Cloaked-2, derived from human and mouse. The cloaked-2
 CC polypeptides can be expressed by standard recombinant methodology. The
 CC cloaked-2 polynucleotides are useful in gene therapy and antisense
 CC therapy. The cloaked-2 polypeptides and polynucleotides are useful for
 CC treating, preventing, ameliorating or detecting diseases and disorders of
 CC the kidney (e.g. anemia, hypertension or low blood pressure), heart (e.g.
 CC cardiac hypertrophy, congestive heart failure, myocardial infarction,
 CC arrhythmias, atherosclerosis, hypertension or low blood pressure),
 CC skeletal muscle (e.g. muscular dystrophy or cachexia), placenta (e.g.
 CC congenital abnormalities or miscarriage), liver (e.g. hepatitis or
 CC cirrhosis), pancreas (e.g. diabetes or pancreatitis), thyroid (e.g.
 CC Grave's disease or myxedema) or adrenal cortex (e.g. Cushing's disease
 CC or Addison's disease), homeostasis or metabolic diseases (e.g. obesity,
 CC cancer or myopathies), infections, or autoimmune diseases. Selective
 CC binding agents may be used to modulate the biological activities of
 CC Cloaked-2 polypeptides or to detect Cloaked-2 polypeptide levels in a
 CC sample. Transgenic non-human animals are useful for drug candidate
 CC screening. The present sequence represents the human cloaked-2
 CC polypeptide sequence.
 XX Sequence 213 AA;
 SQ

CC expressing inactive proteins or to supplement the patients own production
 CC of BEER polypeptides. The nucleic acids may be used for recombinant
 CC production of BEER, gene therapy, antisense therapy, as probes for
 CC diagnostic assays and for functional studies. BEER may be used to raise
 CC antibodies and for identification of BEER modulators. BEER antagonists
 CC may be used to increase bone mineral content for the treatment of
 CC disorders such as osteopenia, osteoporosis, fractures and other
 CC disorders associated with low mineral content.
 XX
 SQ Sequence 213 AA;

Query Match 90.4%; Score 932.5; DB 21; Length 213;
 Best Local Similarity 88.4%; Pred. No. 4.2e-79;
 Matches 168; Conservative 10; Mismatches 7; Indels 5; Gaps 2;
 QY 1 QGWAQFRNDATVPCGLGEYPPPP--ENNQTMNRAENGRRPHHPYDAKDVSEYSREL 58
 Db 24 QGWAQFRNDATVPCGLGEYPPPP--ENNQTMNRAENGRRPHHPYDAKDVSEYSREL 83
 QY 59 HYTRFLTDGPCRSAPVTELVCSGCGPARLLPNAIGRVKWRPNRPGDFRCIPDRYRAQR 118
 Db 84 HFTRYVTDGPCRSAPVTELVCSGCGPARLLPNAIGRVKWRPNRPGDFRCIPDRYRAQR 143
 QY 119 VOLLCPGGAAPRSRKVRVLSVCKKRLTRFHNSQELKDFGTPETARPKGRKPRP---GAK 175
 Db 144 VOLLCPGGAAPRSRKVRVLSVCKKRLTRFHNSQELKDFGTPETARPKGRKPRP---GAK 203
 QY 176 ANQAELENAY 185
 Db 204 ANQAELENAY 213

RESULT 9
 ID AAY97589 standard; Protein; 213 AA.
 XX AAY97589;
 AC AAY97589;
 XX
 DT 05-APR-2001 (first entry)
 XX Human secreted protein PRO7476.
 XX
 KW Secreted protein; human; PRO protein; neoplastic cell growth; tumour;
 KW proliferation; leukaemia; lymphoid malignancy; inflammatory disorder;
 KW angiogenic disorder; immunologic disorder; PRO7476.
 XX
 OS Homo sapiens.
 XX
 PN WO200075317-A2.
 XX

PD 14-DEC-2000.
 XX
 XX 15-MAY-2000; 2000WO-US13358.
 XX
 PR 09-JUN-1999; 99US-0138385.
 PR 20-JUL-1999; 99US-0144790.
 PR 03-AUG-1999; 99US-0146843.
 PR 10-AUG-1999; 99US-0148188.
 PR 17-AUG-1999; 99US-0149320.
 PR 17-AUG-1999; 99US-0149327.
 PR 17-AUG-1999; 99US-0149396.
 PR 20-AUG-1999; 99US-0150114.
 PR 31-AUG-1999; 99US-0151700.
 PR 31-AUG-1999; 99US-0151734.
 XX

(GETH) GENENTECH INC.
 XX
 XX Botstein DA, Goddard A, Gurney AL, Smith V, Watanabe CK, Wood WT;
 PI
 XX WPI; 2001-071075/08.
 DR N-PSDB; AAA91023.
 XX
 PT Antibodies against PRO polypeptides, useful for diagnosing and treating

PT tumours are associated with gene amplification, neoplastic cell growth
 PT and proliferation in mammals -
 XX
 PS Claim 61; Fig 20; 143pp; English.
 XX

CC This sequence is a human PRO protein of the invention. The PRO
 CC proteins are secreted proteins. Antagonists or antibodies of PRO
 CC polypeptides are useful for diagnosing and treating tumours are
 CC associated with gene amplification, neoplastic cell growth and
 CC proliferation in mammals, and those conditions characterised by
 CC overexpression and/or activation of the amplified genes. Such conditions
 CC include benign or malignant tumours (e.g. renal, liver, kidney, bladder,
 CC breast, gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,
 CC thyroid, hepatic carcinomas, sarcomas, glioblastomas and various head and
 CC neck tumours); leukaemias and lymphoid malignancies; neuronal, glial,
 CC astrocytal, hypothalamic, and other glandular, macrophageal, epithelial,
 CC stromal and blastocoealic disorders; and inflammatory, angiogenic and
 CC immunologic disorders. These may further be used to qualitatively or
 CC quantitatively detect the expression of proteins encoded by the
 CC amplified genes, and in tumour diagnostics or prognostics. The PRO
 CC polypeptide or its antagonist may be used for the preparation of a
 CC medicament in the treatment of a condition, which is responsive to the
 CC PRO polypeptide, its antagonist or anti-PRO antibody.
 XX
 SQ Sequence 213 AA;

Query Match 90.4%; Score 932.5; DB 22; Length 213;
 Best Local Similarity 88.4%; Pred. No. 4.2e-79;
 Matches 168; Conservative 10; Mismatches 7; Indels 5; Gaps 2;

QY 1 QGWAQFRNDATVPCGLGEYPPPP--ENNQTMNRAENGRRPHHPYDAKDVSEYSREL 58
 Db 24 QGWAQFRNDATVPCGLGEYPPPP--ENNQTMNRAENGRRPHHPYDAKDVSEYSREL 83
 QY 59 HYTRFLTDGPCRSAPVTELVCSGCGPARLLPNAIGRVKWRPNRPGDFRCIPDRYRAQR 118
 Db 84 HFTRYVTDGPCRSAPVTELVCSGCGPARLLPNAIGRVKWRPNRPGDFRCIPDRYRAQR 143
 QY 119 VOLLCPGGAAPRSRKVRVLSVCKKRLTRFHNSQELKDFGTPETARPKGRKPRP---GAK 175
 Db 144 VOLLCPGGAAPRSRKVRVLSVCKKRLTRFHNSQELKDFGTPETARPKGRKPRP---GAK 203
 QY 176 ANQAELENAY 185
 Db 204 ANQAELENAY 213

RESULT 10
 ID ABG34061 standard; Protein; 213 AA.
 XX ABG34061;
 AC ABG34061;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Human Pro peptide #32.
 XX
 KW Human; PRO; secreted protein; transmembrane protein;
 KW genetic disorder; tumour; cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO200224888-A2.
 XX
 PD 28-MAR-2002.
 XX
 XX 29-AUG-2001; 2001WO-US27099.
 PF
 XX 01-SEP-2000; 2000US-229896P.
 PR 05-SEP-2000; 2000US-230621P.
 PR 22-SEP-2000; 2000US-235147P.
 PR 10-NOV-2000; 2000WO-US30873.
 PR 12-JAN-2001; 2001US-261878P.

CC Grave's disease or myxedema) or adrenal cortex (e.g. Cushing's disease
 CC or Addison's disease), homeostasis or metabolic diseases (e.g. obesity,
 CC cancer or myopathies), infections, or autoimmune diseases. Selective
 CC binding agents may be used to modulate the biological activities of
 CC Cloaked-2 polypeptides or to detect Cloaked-2 polypeptide levels in a
 CC sample. Transgenic non-human animals are useful for drug candidate
 CC screening. The present sequence represents the human cloaked-2
 CC mature polypeptide sequence.

XX Sequence 190 AA;
 SQ Query Match 90.4%; Score 932.5; DB 23; Length 190;
 Best Local Similarity 88.4%; Pred. No. 3.7e-79; Indels 5; Gaps 2;
 Matches 168; Conservative 10; Mismatches 7;

QY 1 QGWAQFNDAEIVPGIAGEYEPPEPP--ENNQTMNRAENGRRPHHPHYDAKDVSEYSCREL 58
 DB 1 QGWAQFNDAEIVPGIAGEYEPPEPP--ENNQTMNRAENGRRPHHPHYDAKDVSEYSCREL 60
 QY 59 HYTRFLTDGPCRSAPVTELVCSGQCGPARLLPNAIGRVKWRPNPGRPDRCIPDRYRAQR 118
 DB 61 HFTYVTDGPCRSAPVTELVCSGQCGPARLLPNAIGRVKWRPNPGRPDRCIPDRYRAQR 120
 QY 119 VQLLCPGGAAPRSKRVLVASCKCKRLTRFHNQSELKDFGTEAARPKGKRP---GAK 175
 DB 121 VQLLCPGGAAPRSKRVLVASCKCKRLTRFHNQSELKDFGTEAARPKGKRP---GAK 180
 QY 176 ANQAELENAY 185
 DB 181 ANQAELENAY 190

RESULT 7
 AAB26106
 ID AAB26106 standard; Protein; 213 AA.
 XX AC AAB26106;

15-JAN-2001 (first entry)

Human DAN/Cerberus-related protein 6 (hDCR6) (exons 1 and 4).

Human; DNA/Cerberus-related protein 6; hDCR6; morphogenic protein;
 antagonist; BMP; cell growth; cell differentiation; bone formation;
 gene therapy.

XX Homo sapiens.

XX WO200055193-A2.

XX 21-SEP-2000.

XX 02-MAR-2000; 2000WO-US05537.

XX 12-MAR-1999; 99US-0124118.

XX (REG-) REGENERON PHARM INC.

XX Economides AN;

XX WPI; 2000-638179/61.

XX N-PSDB; AAA94051.

XX Novel isolated, human DNA/Cerberus related protein 6 which include
 PT natural homologue, and polypeptides comprising DCR6 domain and nucleic
 PT acids encoding the proteins which are useful as probes and primers -

XX Claim 8; Fig 3; 40pp; English.

XX The present sequence comprises the amino acid sequence encoded by exons 1
 CC and 4 of the human DAN/Cerberus-related protein 6 (hDCR6) coding
 CC sequence. The coding sequence was isolated from a human kidney cDNA
 CC library containing exons 1 and 4 of the sequence. hDCR6 is closely

CC related to the DAN and DCR5 proteins, both of which act as antagonists of
 CC morphogenic proteins such as BMP. It is possible that the hDCR6 gene and
 CC protein can be used as immunogens, modulators of cell function, growth
 CC and differentiation, to reduce undesirable bone formation, to identify
 CC DCR6 binding agents, in diagnosis, and in gene therapy.

XX Sequence 213 AA;

XX Query Match 90.4%; Score 932.5; DB 21; Length 213;
 Best Local Similarity 88.4%; Pred. No. 4.2e-79;
 Matches 168; Conservative 10; Mismatches 7; Indels 5; Gaps 2;

QY 1 QGWAQFNDAEIVPGIAGEYEPPEPP--ENNQTMNRAENGRRPHHPHYDAKDVSEYSCREL 58
 DB 24 QGWAQFNDAEIVPGIAGEYEPPEPP--ENNQTMNRAENGRRPHHPHYDAKDVSEYSCREL 83
 QY 59 HYTRFLTDGPCRSAPVTELVCSGQCGPARLLPNAIGRVKWRPNPGRPDRCIPDRYRAQR 118
 DB 84 HFTYVTDGPCRSAPVTELVCSGQCGPARLLPNAIGRVKWRPNPGRPDRCIPDRYRAQR 143
 QY 119 VQLLCPGGAAPRSKRVLVASCKCKRLTRFHNQSELKDFGTEAARPKGKRP---GAK 175
 DB 144 VQLLCPGGAAPRSKRVLVASCKCKRLTRFHNQSELKDFGTEAARPKGKRP---GAK 203
 QY 176 ANQAELENAY 185
 DB 204 ANQAELENAY 213

RESULT 8
 AAY96429
 ID AAY96429 standard; Protein; 213 AA.
 XX AC AAY96429;

12-SEP-2000 (first entry)

Human TGF-beta binding protein (BEER).

osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 BEER; gene therapy; antisense therapy; fracture; bone mineralization.

XX Homo sapiens.

XX WO200032773-A1.

XX 08-JUN-2000.

XX 24-NOV-1999; 99WO-US27990.

XX 27-NOV-1998; 98US-0110283.

XX (DARW-) DARWIN DISCOVERY LTD.

XX Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepers BW;

XX Van Ness J, Winkler DG;

XX WPI; 2000-412321/35.

XX N-PSDB; AAA29055.

XX Nucleic acids (1) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures

XX Claim 2; Page 116; 162pp; English.

XX This shows the human transforming growth factor-beta (TGF-beta)
 CC binding protein designated hBEER. The cDNA and protein may be used for
 CC prevention, treatment and diagnosis of diseases associated with
 CC inappropriate BEER expression. For example, they may be used to treat
 CC disorders associated with decreased TGF-beta BP expression. The cDNA or
 CC vectors may be administered to treat diseases by rectifying mutations or
 CC deletions in a patient's genome that affect the activity of BEER by

Db 84 HYTRFVTDGFCRSAPKPVTELVCSCGCGPARLLPNAIGRKWWRNGPDPFCIPDRYRAQR 143
 QY 119 VOLLCPGGAAPSRKRVRLVASCCKRLTRFNQSELKDFGPETARPOKGRKPRP---GAK 175
 Db 144 VOLLCPGGAAPSRKRVRLVASCCKRLTRFNQSELKDFGPETARPOKGRKPRPARGAK 203
 QY 176 ANQAELENAY 185
 Db 204 ANQAELENAY 213

RESULT 5
 AAY96431
 ID AAY96431 standard; Protein; 213 AA.
 XX
 AC AAY96431;
 DT 12-SEP-2000 (first entry)
 DE Vervet TGF-beta binding protein (BEER).
 KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KW BEER; gene therapy; antisense therapy; fracture; bone mineralization.
 XX Cercopithecus pygerythrus.
 OS
 PN W0200032773-A1.
 PD 08-JUN-2000.
 XX
 PF 24-NOV-1999; 99WO-US27990.
 XX
 PR 27-NOV-1998; 98US-0110283.
 XX
 PA (DARW-) DARWIN DISCOVERY LTD.
 PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepers BW;
 PI Van Ness J, Winkler DG;
 XX
 DR WPI; 2000-412321/35.
 DR N-PSDB; AAA29057.
 XX

Nucleic acids (1) encoding a transforming growth factor beta binding protein, useful for identifying agents for treating osteopenia, osteoporosis and fractures
 Claim 4; Page 122-123; 162pp; English.

This shows a vervet transforming growth factor-beta (TGF-beta) binding protein designated vBEER. The cDNA and protein may be used for prevention, treatment and diagnosis of diseases associated with inappropriate BEER expression. For example, they may be used to treat disorders associated with decreased TGF-beta BP expression. The cDNA or vectors may be administered to treat diseases by rectifying mutations or deletions in a patient's genome that affect the activity of BEER by expressing inactive proteins or to supplement the patients own production of BEER polypeptides. The nucleic acids may be used for recombinant production of BEER, gene therapy, antisense therapy, as probes for diagnostic assays and for functional studies. BEER may be used to raise antibodies and for identification of BEER modulators. BEER antagonists may be used to increase bone mineral content for the treatment of disorders such as osteopenia, osteoporosis, fractures and other disorders associated with low mineral content.

XX Sequence 213 AA;
 SQ

Query Match
 Best Local Similarity 92.2%; Score 951.5; DB 21; Length 213;
 Matches 171; Conservative 10; Mismatches 4; Indels 5; Gaps 2;
 QY 1 QGWAQFNDATVPIGLGYPPEPPP--ENNQTMNRAENGRRPHHPYDAKDVSYSREL 58
 Db 24 QGWAQFNDATVPIGLGYPPEPPP--ENNQTMNRAENGRRPHHPYDAKDVSYSREL 58

QY 59 HYTRFVTDGFCRSAPKPVTELVCSCGCGPARLLPNAIGRKWWRNGPDPFCIPDRYRAQR 118
 Db 84 HFRYVTDGFCRSAPKPVTELVCSCGCGPARLLPNAIGRKWWRNGPDPFCIPDRYRAQR 143
 QY 119 VOLLCPGGAAPSRKRVRLVASCCKRLTRFNQSELKDFGPETARPOKGRKPRP---GAK 175
 Db 144 VOLLCPGGAAPSRKRVRLVASCCKRLTRFNQSELKDFGPETARPOKGRKPRPARGAK 203
 QY 176 ANQAELENAY 185
 Db 204 ANQAELENAY 213

RESULT 6
 ABB07207
 ID ABB07207 standard; Protein; 190 AA.
 XX
 AC ABB07207;
 DT 26-MAR-2002 (first entry)
 DE Human cloaked-2 polypeptide mature protein sequence.
 XX
 KW Cloaked-2; cysteine knot motif; nephrotropic; cardiant; immunomodulator;
 KW hepatotropic; antiinflammatory; antithyroid; cytostatic; neuroprotective;
 KW antianemic; hypotensive; antiarrhythmic; antiarteriosclerotic; muscular;
 KW antidiabetic; anorectic; gene therapy; cell therapy; antisense therapy;
 XX human.
 OS Homo sapiens.
 FH
 FT Key Location/Qualifiers
 FT Misc-difference 1..16 /note= "this region has been repeated twice in the
 FT sequence provided in the sequence listing but has
 FT been indicated correctly in the sequence in the
 FT Figure"
 XX
 PN W0200192308-A2.
 XX
 PD 06-DEC-2001.
 XX
 PF 29-MAY-2001; 2001WO-US17478.
 XX
 PR 01-JUN-2000; 2000US-208550P.
 PR 04-AUG-2000; 2000US-223542P.
 XX
 PA (AMGE-) AMGEN INC.
 XX
 PI Paszty CJ, Gao Y;
 XX
 DR WPI; 2002-114325/15.
 DR N-PSDB; ABA94293.
 XX

New human and mouse cysteine-knot polypeptide designated as Cloaked-2, for treating or preventing kidney, heart (e.g. myocardial infarction) or liver (e.g. hepatitis) diseases
 Claim 13; Fig 1; 170pp; English.

The invention relates to polypeptides comprising a cysteine knot motif and designated as Cloaked-2, derived from human and mouse. The cloaked-2 polypeptides can be expressed by standard recombinant methodology. The cloaked-2 polynucleotides are useful in gene therapy and antisense therapy. The cloaked-2 polypeptides and polynucleotides are useful for treating, preventing, ameliorating or detecting diseases and disorders of the kidney (e.g. anemia, hypertension or low blood pressure), heart (e.g. cardiac hypertrophy, congestive heart failure, myocardial infarction, arrhythmias, atherosclerosis, hypertension or low blood pressure), skeletal muscle (e.g. muscular dystrophy or cachexia), placenta (e.g. congenital abnormalities or miscarriage), liver (e.g. hepatitis or cirrhosis), pancreas (e.g. diabetes or pancreatitis), thyroid (e.g.

antidiabetic; anorectic; gene therapy; cell therapy; antisense therapy; mouse.

Mus musculus.

Key Location/Qualifiers
 Peptide 1..23 /note= "signal peptide"
 Protein 24 /note= "mature protein (AB070208)"
 Misc-difference 198..200
 /note= "this region is missing in the sequence provided in the sequence listing but has been indicated correctly in the sequence in the Figure"

WO200192308-A2.
 06-DEC-2001.
 29-MAY-2001; 2001WO-US17478.
 01-JUN-2000; 2000US-208550P.
 04-AUG-2000; 2000US-223542P.
 (AMGE-) AMGEN INC.
 Paszty CJ, Gao Y;
 WPI; 2002-114325/15.
 N-PSDB; ABA94294.
 New human and mouse cystine-knot polypeptide designated as Cloaked-2, for treating or preventing kidney, heart (e.g. myocardial infarction) or liver (e.g. hepatitis) diseases -
 Example 2; Fig 2; 170pp; English.

The invention relates to polypeptides comprising a cysteine knot motif and designated as Cloaked-2, derived from human and mouse. The cloaked-2 polypeptides can be expressed by standard recombinant methodology. The cloaked-2 polynucleotides are useful in gene therapy and antisense therapy. The cloaked-2 polypeptides and polynucleotides are useful for treating, preventing, ameliorating or detecting diseases and disorders of the kidney (e.g. anemia, hypertension or low blood pressure), heart (e.g. cardiac hypertrophy, congestive heart failure, myocardial infarction, arrhythmias, atherosclerosis, hypertension or low blood pressure), skeletal muscle (e.g. muscular dystrophy or cachexia), placenta (e.g. cirrhosis), pancreas (e.g. diabetes or pancreatitis), liver (e.g. hepatitis or Grave's disease or myxedema) or adrenal cortex (e.g. Cushing's disease or Addison's disease), infections, or autoimmune diseases. Selective binding agents may be used to modulate the biological activities of Cloaked-2 polypeptides or to detect Cloaked-2 polypeptide levels in a sample. Transgenic non-human animals are useful for drug candidate screening. The present sequence represents the mouse cloaked-2 polypeptide.

Sequence 211 AA;
 Query Match 98.9%; Score 1020.5; DB 23; Length 211;
 Best Local Similarity 98.4%; Pred. No. 2.5e-87;
 Matches 185; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

1 QGQAFRNDATVIFGLGEYPPPPNNQTNRAENGRRPPHPHYDAKDVSEYSCRELHY 60
 24 QGQAFRNDATVIFGLGEYPPPPNNQTNRAENGRRPPHPHYDAKDVSEYSCRELHY 83
 61 TRFLTDGCRSAKPVTELVCSGCGPARLLPNAIGRVKWRPNPDPFCIPDRYRAQRVQ 120
 84 TRFLTDGCRSAKPVTELVCSGCGPARLLPNAIGRVKWRPNPDPFCIPDRYRAQRVQ 143
 121 LLCFGGAAPRSRKVLVASCCKRLTRFHNSQLKDFGPETARPKGRKPRPGARGKAN 177

Db 144 LLCFGGAAPRSRKVLVASCCKRLTRFHNSQLKDFGPETARPKGRKPRPGARGKAN 203
 Qy 178 QAELENAY 185
 Db 204 QAELENAY 211

RESULT 4
 AAY96433
 ID AAY96433 standard; Protein; 213 AA.
 AC AAY96433;
 DT 12-SEP-2000 (first entry)
 DE Rat TGF-beta binding protein (BEER).
 KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein; BEER; gene therapy; antisense therapy; fracture; bone mineralization.
 OS Rattus norvegicus.
 PN WO200032773-A1.
 PD 08-JUN-2000.
 PF 24-NOV-1999; 99WO-US27990.
 PR 27-NOV-1998; 98US-0110283.
 PA (DARW-) DARWIN DISCOVERY LTD.
 PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepers BW;
 PI Van Ness J, Winkler DG;
 WPI; 2000-412321/35.
 N-PSDB; AAY29059.

Nucleic acids (I) encoding a transforming growth factor beta binding protein, useful for identifying agents for treating osteoporosis, osteoporosis and fractures
 Claim 6; Page 125-126; 162pp; English.

This shows a rat transforming growth factor-beta (TGF-beta) binding protein designated rBEER. The CDNA and protein may be used for prevention, treatment and diagnosis of diseases associated with inappropriate BEER expression. For example, they may be used to treat disorders associated with decreased TGF-beta BP expression. The CDNA or vectors may be administered to treat diseases by rectifying mutations or deletions in a patient's genome that affect the activity of BEER by expressing inactive proteins or to supplement the patient's own production of BEER polypeptides. The nucleic acids may be used for recombinant production of BEER, gene therapy, antisense therapy, as probes for diagnostic assays and for functional studies. BEER may be used to raise antibodies and for identification of BEER modulators. BEER antagonists may be used to increase bone mineral content for the treatment of disorders such as osteoporosis, osteoporosis, fractures and other disorders associated with low mineral content.

Sequence 213 AA;
 Query Match 95.2%; Score 982.5; DB 21; Length 213;
 Best Local Similarity 94.2%; Pred. No. 9e-84; 3; Indels 5; Gaps 2;
 Matches 179; Conservative 3; Mismatches 3

1 QGQAFRNDATVIFGLGEYPPPP--ENNQTMNRAENGRRPPHPHYDAKDVSEYSCREL 58
 24 QGQAFRNDATVIFGLGEYPPPPQLENNQTMNRAENGRRPPHPHYDTKDVSEYSCREL 83
 59 HYTRFLTDGCRSAKPVTELVCSGCGPARLLPNAIGRVKWRPNPDPFCIPDRYRAQR 118

(AMGE-) AMGEN INC.
 Paszty CJ, Gao Y;
 WPI: 2002-114325/15.
 N-PSDB; ABA94294.
 New human and mouse cysteine-knot polypeptide designated as Cloaked-2,
 PT for treating or preventing kidney, heart (e.g. myocardial infarction)
 or liver (e.g. hepatitis) diseases -
 XX
 PS Claim 13; Fig 2; 170pp; English.
 XX
 CC The invention relates to polypeptides comprising a cysteine knot motif
 CC and designated as Cloaked-2, derived from human and mouse. The cloaked-2
 CC polypeptides can be expressed by standard recombinant methodology. The
 CC cloaked-2 polynucleotides are useful in gene therapy and antisense
 CC therapy. The cloaked-2 polypeptides and polynucleotides are useful for
 CC treating, preventing, ameliorating or detecting diseases and disorders of
 CC the kidney (e.g. anemia, hypertension or low blood pressure), heart (e.g.
 CC cardiac hypertrophy, congestive heart failure, myocardial infarction,
 CC arrhythmias, atherosclerosis, hypertension or low blood pressure),
 CC skeletal muscle (e.g. muscular dystrophy or cachexia), placenta (e.g.
 CC congenital abnormalities or miscarriage), liver (e.g. hepatitis or
 CC cirrhosis), pancreas (e.g. diabetes or pancreatitis), thyroid (e.g.
 CC Grave's disease or myxedema) or adrenal cortex (e.g. Cushing's disease
 CC or Addison's disease), homeostasis or metabolic diseases (e.g. obesity,
 CC cancer or myopathies), infections, or autoimmune diseases. Selective
 CC binding agents may be used to modulate the biological activities of
 CC Cloaked-2 polypeptides or to detect Cloaked-2 polypeptide levels in a
 CC sample. Transgenic non-human animals are useful for drug candidate
 CC screening. The present sequence represents the mouse cloaked-2 mature
 CC polypeptide.
 XX
 SQ Sequence 188 AA;
 Query Match 98.9%; Score 1020.5; DB 23; Length 188;
 Best Local Similarity 98.4%; Pred. No. 2.2e-87;
 Matches 185; Conservative 0; Mismatches 0; Indels 3; Gaps 1;
 QY 1 QGWAQFRNDATVPIGLGEYEPPEPPNNQTMRAENGRRPHHPYDAKDVSEYSCRELHY 60
 DB 1 QGWAQFRNDATVPIGLGEYEPPEPPNNQTMRAENGRRPHHPYDAKDVSEYSCRELHY 60
 QY 61 TRFTLDGPCRSKAPVTELVCSGCGCPARLLPNAIGRVKWRPNPGRFCIPDRYRAQRVQ 120
 DB 61 TRFTLDGPCRSKAPVTELVCSGCGCPARLLPNAIGRVKWRPNPGRFCIPDRYRAQRVQ 120
 QY 121 LLCFGGAAPRSRKVRLVASCKCKRLTRFHNSQSELKDFGPETARPOKGRKPRP---GAKAN 177
 DB 121 LLCFGGAAPRSRKVRLVASCKCKRLTRFHNSQSELKDFGPETARPOKGRKPRP---GAKAN 177
 QY 178 QAELENAY 185
 DB 181 QAELENAY 188
 RESULT 2
 AAY96432
 ID AAY96432 standard; Protein; 211 AA.
 AC AAY96432;
 XX
 XX 12-SEP-2000 (first entry)
 DT
 DE Murine TGF-beta binding protein (BEER).
 XX osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KW BEER; gene therapy; antisense therapy; fracture; bone mineralization.
 XX
 OS Mus musculus.
 XX
 PN WO200032773-A1.

XX
 PD 08-JUN-2000.
 XX
 PF 24-NOV-1999; 99WO-US27990.
 XX
 PR 27-NOV-1998; 98US-0110283.
 XX
 PA (DARW-) DARWIN DISCOVERY LTD.
 XX
 PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepers BW;
 PI Van Ness J, Winkler DG;
 XX
 DR WPI: 2000-412321/35.
 DR N-PSDB; AAA29058.
 XX
 PT Nucleic acids (1) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures
 PS Claim 5; Page 124; 162pp; English.
 XX
 CC This shows a murine transforming growth factor-beta (TGF-beta)
 CC binding protein designated mBEER. The cDNA and protein may be used for
 CC prevention, treatment and diagnosis of diseases associated with
 CC inappropriate BEER expression. For example, they may be used to treat
 CC disorders associated with decreased TGF-beta BP expression. The cDNA or
 CC vectors may be administered to treat diseases by rectifying mutations or
 CC deletions in a patient's genome that affect the activity of BEER by
 CC expressing inactive proteins or to supplement the patient's own production
 CC of BEER polypeptides. The nucleic acids may be used for recombinant
 CC production of BEER, gene therapy, antisense therapy, as probes for
 CC diagnostic assays and for functional studies. BEER may be used to raise
 CC antibodies and for identification of BEER modulators. BEER antagonists
 CC may be used to increase bone mineral content for the treatment of
 CC disorders such as osteopenia, osteoporosis, fractures and other disorders
 CC associated with low mineral content.
 XX
 SQ Sequence 211 AA;
 Query Match 98.9%; Score 1020.5; DB 21; Length 211;
 Best Local Similarity 98.4%; Pred. No. 2.5e-87;
 Matches 185; Conservative 0; Mismatches 0; Indels 3; Gaps 1;
 QY 1 QGWAQFRNDATVPIGLGEYEPPEPPNNQTMRAENGRRPHHPYDAKDVSEYSCRELHY 60
 DB 24 QGWAQFRNDATVPIGLGEYEPPEPPNNQTMRAENGRRPHHPYDAKDVSEYSCRELHY 83
 QY 61 TRFTLDGPCRSKAPVTELVCSGCGCPARLLPNAIGRVKWRPNPGRFCIPDRYRAQRVQ 120
 DB 84 TRFTLDGPCRSKAPVTELVCSGCGCPARLLPNAIGRVKWRPNPGRFCIPDRYRAQRVQ 143
 QY 121 LLCFGGAAPRSRKVRLVASCKCKRLTRFHNSQSELKDFGPETARPOKGRKPRP---GAKAN 177
 DB 144 LLCFGGAAPRSRKVRLVASCKCKRLTRFHNSQSELKDFGPETARPOKGRKPRP---GAKAN 203
 QY 178 QAELENAY 185
 DB 204 QAELENAY 211
 RESULT 3
 ABB07210
 ID ABB07210 standard; Protein; 211 AA.
 XX
 AC ABB07210;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Mouse cloaked-2 polypeptide sequence.
 XX
 KW Cloaked-2; cysteine knot motif; nephrotropic; cardiant; immunomodulator;
 KW hepatotropic; antiinflammatory; antithyroid; cytoprotective;
 KW anianemic; hypotensive; antiarrhythmic; antiarteriosclerotic; muscular;

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:06:06 ; Search time 51.8 Seconds
(without alignments)
475.895 Million cell updates/sec

Title: US-09-867-274-4
Perfect score: 1032
Sequence: 1 OGMQAFRNDATVIFGLGEY.....KGRKPRFCANQALENAY 185

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 101002:*

1: /SID22/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*

2: /SID22/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*

3: /SID22/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*

4: /SID22/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*

5: /SID22/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*

6: /SID22/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*

7: /SID22/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*

8: /SID22/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*

9: /SID22/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*

10: /SID22/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*

11: /SID22/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*

12: /SID22/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*

13: /SID22/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*

14: /SID22/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*

15: /SID22/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*

16: /SID22/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*

17: /SID22/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*

18: /SID22/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*

19: /SID22/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*

20: /SID22/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*

21: /SID22/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*

22: /SID22/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

23: /SID22/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1020.5	98.9	188	23	Mouse cloaked-2 po
2	1020.5	98.9	211	21	Murine TGF-beta bi
3	1020.5	98.9	211	23	Mouse cloaked-2 po
4	982.5	95.2	213	21	Rat TGF-beta bindi
5	951.5	92.2	213	21	Vervet TGF-beta bi
6	932.5	90.4	190	23	Human cloaked-2 po
7	932.5	90.4	213	21	Human DAN/Cerberus
8	932.5	90.4	213	21	Human TGF-beta bin
9	932.5	90.4	213	22	Human secreted pro
10	932.5	90.4	213	23	Human Pro peptide

11	932.5	90.4	213	23	ABB07209	Human cloaked-2 po
12	932.5	90.4	213	23	AAE17089	Human osteolevin p
13	923.5	89.5	213	21	AAV96430	Human TGF-beta bin
14	923.5	89.5	213	21	AAV96436	Human TGF-beta bi
15	839	81.3	176	21	AAV96434	Bovine TGF-beta bi
16	837.5	81.2	367	21	AAE26105	Human DAN/Cerberus
17	687.5	66.6	139	22	AAW73067	Human bone marrow
18	687.5	66.6	139	23	ABG42911	Human peptide enco
19	330	32.0	206	21	AAV75981	Murine skin cell p
20	330	32.0	206	21	AAV76031	Murine skin cell p
21	330	32.0	206	22	AAE55920	Skin cell protein,
22	330	32.0	206	22	AAE55970	Skin cell protein,
23	330	32.0	206	22	ABE72120	Murine protein iso
24	330	32.0	206	23	ABE72170	Murine protein iso
25	328	31.8	206	23	AAW09408	Human small CCN-li
26	328	31.8	206	23	AAW58704	Human small CCN-li
27	328	31.8	206	20	AAW5711	Homo sapiens fetal
28	328	31.8	206	21	AAE10233	Human adult retina
29	328	31.8	206	23	ABE95457	Human angiogenesis
30	328	31.8	206	23	ABE95200	Human small CCN-li
31	328	31.8	206	23	ABE95210	SCGF CNN family pr
32	328	31.8	206	23	ABE94851	Human PRO532 prote
33	324	31.4	183	23	ABE07213	Human cloaked-1 po
34	324	31.4	206	23	AAU82653	Human breast tumou
35	312	30.2	182	23	AAU82655	Human breast tumou
36	137	13.3	50	20	AAV12009	Human 5' EST seque
37	109.5	10.6	116	18	AAW27654	Secreted protein A
38	109.5	10.6	116	18	AAW44090	Human secreted pro
39	98.5	9.5	102	20	AAV03225	Amino acid sequenc
40	98.5	9.5	102	21	AAE10277	Human fetal kidney
41	90.5	8.8	184	19	AAW29727	DAN and b57 protei
42	90	8.7	672	22	AAE3107	Human polypeptide
43	90	8.7	712	21	AAE43147	Human ORFX ORF2911
44	90	8.7	739	22	AAE39106	Human polypeptide
45	88.5	8.6	182	19	AAW29726	DAN and b57 protei

ALIGNMENTS

RESULT 1

ABB07208

ID ABB07208 standard; Protein; 188 AA.

XX

AC ABB07208;

XX

XX 26-MAR-2002 (first entry)

DT Mouse cloaked-2 polypeptide mature protein sequence.

DE Cloaked-2; cysteine knot motif; nephrotropic; cardiant; immunomodulator;

XX hepatotropic; antiinflammatory; antithyroid; cytosolic; neuroprotective;

KW antianemic; hypotensive; antiarrhythmic; antiarteriosclerotic; muscular;

KW antididiabetic; anorectic; gene therapy; cell therapy; antisense therapy;

KW mouse.

XX Mus musculus.

OS

XX

FT Key Location/Qualifiers

FT Misc-difference 175..177

FT /note= "this region is missing in the sequence provided

FT in the sequence listing but has been indicated

FT correctly in the sequence in the Figure"

XX WO200192308-A2.

XX

XX 06-DEC-2001.

XX

XX 29-MAY-2001; 2001WO-US17478.

XX

XX 01-JUN-2000; 2000US-208550P.

PR

PR 04-AUG-2000; 2000US-223542F.

XX

; TITLE OF INVENTION: Interferon-Suppressing Placental Lactogen Peptides
; FILE REFERENCE: 1101

; CURRENT APPLICATION NUMBER: US/09/876,478
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: 60/210,082
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: MS-DOS
; SEQ ID NO 16
; LENGTH: 56
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
; FEATURE:
; OTHER INFORMATION: N-TERMINAL 56 RESIDUES OF HPRL
US-09-876-478-16

Query Match 3.2%; Score 6; DB 10; Length 56;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 123 CPGGAA 128
Db 32 CPGGAA 37
|||||

RESULT 14
US-09-510-332-72
; Sequence 72, Application US/09510332
; Publication No. US2003002278A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Ryba, Nick
; APPLICANT: Mueller, Ken
; APPLICANT: Hoon, Mark
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: T2R, a No. US2003002278A1el Family of Taste Receptors
; FILE REFERENCE: 02307E-098010US
; CURRENT APPLICATION NUMBER: US/09/510,332
; CURRENT FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 09/393,634
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 72
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human T2R46 (hGR46)
US-09-510-332-72

Query Match 3.2%; Score 6; DB 9; Length 59;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 77 ELVCSG 82
Db 54 ELVCSG 59
|||||

RESULT 15
US-09-864-761-35511
; Sequence 35511, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1

; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 35511
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AJ239318.3
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.96
; OTHER INFORMATION: EST HUMAN HIT: AW962016.1, EVALUE 2.00e-20
; OTHER INFORMATION: SWISSPROT HIT: P41997, EVALUE 3.00e-05
US-09-864-761-35511

Query Match 3.2%; Score 6; DB 10; Length 60;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 157 FGPETA 162
Db 20 FGPETA 25
|||||

Search completed: March 28, 2003, 14:24:58
Job time : 12.84 secs

us-09-867-274-4.oligo.rapb

Mon Mar 31 09:44:33 2003

```
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34998
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007898.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN HELLO, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2
; NAME/KEY: unsure
; LOCATION: 4
; NAME/KEY: unsure
; LOCATION: 39
; US-09-864-761-34998

Query Match 3.2%; Score 6; DB 10; Length 41;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 PETARP 164
Db 5 PETARP 10

RESULT 11
US-09-925-297-836
; Sequence 836, Application US/09925297
; Patent No. US20020081659A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PAL05
; CURRENT APPLICATION NUMBER: US/09/925,297
; CURRENT FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: PCT/US00/05989
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 928
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 836
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (4)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (6)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (28)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (42)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-925-297-836

Query Match 3.2%; Score 6; DB 10; Length 46;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 73 KPVTTEL 78
Db 34 KPVTTEL 39

RESULT 12
US-10-001-857-164
; Sequence 164, Application US/10001857
; Publication No. US20020183500A1
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Hervé
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prote
; FILE REFERENCE: DEX-0273
; CURRENT APPLICATION NUMBER: US/10/001,857
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/252,054
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 164
; LENGTH: 53
; TYPE: PRT
; ORGANISM: Homo sapien
; US-10-001-857-164

Query Match 3.2%; Score 6; DB 9; Length 53;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 GGRPPH 42
Db 21 GGRPPH 26

RESULT 13
US-09-876-478-16
; Sequence 16, Application US/09876478
; Patent No. US20020032154A1
; GENERAL INFORMATION:
; APPLICANT: Peyman, John A.
```



```

; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 44917
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC021000.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.76
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.95
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.86
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.93
US-09-864-761-44917

```

```

Query Match 3.2%; Score 6; DB 10; Length 26;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 171 RPKAKA 176
Db 7 RPKAKA 12

```

```

RESULT 9
US-09-864-761-40983
; Sequence 40983, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Acomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30

```

```

; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 40983
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC018742.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.2
; OTHER INFORMATION: SWISSPROT HIT: P22297, EVALUE 2.20e+00
; OTHER INFORMATION: EST_HUMAN HIT: A1220244.1, EVALUE 6.00e-03
US-09-864-761-40983

```

```

Query Match 3.2%; Score 6; DB 10; Length 40;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 25 PENNOT 30
Db 24 PENNOT 29

```

```

RESULT 10
US-09-864-761-34998
; Sequence 34998, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Acomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03

```

us-09-867-274-4.oligo.rapb

Mon Mar 31 09:44:33 2003

```

QY 65 TDGCRSAKPVTELVCSGCGPARLLPNAIGR 96
      |||||||
Db 67 TDGCRSAKPVTELVCSGCGPARLLPNAIGR 98
      |||||||

RESULT 5
US-09-867-274-5
; Sequence 5, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 5
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-867-274-5

Query Match 17.3%; Score 32; DB 10; Length 213;
Best Local Similarity 100.0%; Pred. No. 2.4e-24;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 TDGCRSAKPVTELVCSGCGPARLLPNAIGR 96
      |||||||
Db 90 TDGCRSAKPVTELVCSGCGPARLLPNAIGR 121
      |||||||

RESULT 6
US-09-732-680A-2
; Sequence 2, Application US/09732680A
; Patent No. US20020046412A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Aktiengesellschaft
; TITLE OF INVENTION: Nucleic acids coding for new acetylcholine receptor beta subunits
; TITLE OF INVENTION: insects
; FILE REFERENCE: Le A 34 147
; CURRENT APPLICATION NUMBER: US/09/732,680A
; CURRENT FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: DE 199 59 582.8
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 441
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-732-680A-2

Query Match 3.8%; Score 7; DB 10; Length 441;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 PGAKANQ 178
      |||||||
Db 404 PGAKANQ 410
      |||||||

RESULT 7
US-09-815-242-13801
; Sequence 13801, Application US/09815242
; Patent No. US20020061589A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13801
; LENGTH: 535
; TYPE: PRT
; ORGANISM: Salmonella typhi
; NAME/KEY: VARIANT
; LOCATION: (1)...(535)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-815-242-13801

Query Match 3.8%; Score 7; DB 10; Length 535;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 TEVIPGL 17
      |||||||
Db 72 TEVIPGL 78
      |||||||

RESULT 8
US-09-864-761-44917
; Sequence 44917, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30

```

```
RESULT 2
US-09-867-274-6
; Sequence 6, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; APPLICANT: Gao, Yongming
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-867-274-6

Query Match      100.0%; Score 185; DB 10; Length 208;
Best Local Similarity 100.0%; Pred. No. 2e-177;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QGWAFRNDATVPIGLGEYEPPEPPENNTMNAENGRRPHHPYDAKDVSEYSCRELHY 60
Db 24 QGWAFRNDATVPIGLGEYEPPEPPENNTMNAENGRRPHHPYDAKDVSEYSCRELHY 83
Qy 61 TRFTDTGPCRSKPVTELVCSGCGPARLLPNAIGRVKWRPNPDPFCIPDRIYRAQRVQ 120
Db 84 TRFTDTGPCRSKPVTELVCSGCGPARLLPNAIGRVKWRPNPDPFCIPDRIYRAQRVQ 143
Qy 121 LLCFGAAPSRSKRVLRVASCCKRLTRFHNQSELKDFGPTETAPQKGRPRFGKANKQAE 180
Db 144 LLCFGAAPSRSKRVLRVASCCKRLTRFHNQSELKDFGPTETAPQKGRPRFGKANKQAE 203
Qy 181 LENAY 185
Db 204 LENAY 208

RESULT 3
US-09-864-761-47109
; Sequence 47109, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667

Query Match      100.0%; Score 185; DB 10; Length 208;
Best Local Similarity 100.0%; Pred. No. 2e-177;
Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QGWAFRNDATVPIGLGEYEPPEPPENNTMNAENGRRPHHPYDAKDVSEYSCRELHY 60
Db 24 QGWAFRNDATVPIGLGEYEPPEPPENNTMNAENGRRPHHPYDAKDVSEYSCRELHY 83
Qy 61 TRFTDTGPCRSKPVTELVCSGCGPARLLPNAIGRVKWRPNPDPFCIPDRIYRAQRVQ 120
Db 84 TRFTDTGPCRSKPVTELVCSGCGPARLLPNAIGRVKWRPNPDPFCIPDRIYRAQRVQ 143
Qy 121 LLCFGAAPSRSKRVLRVASCCKRLTRFHNQSELKDFGPTETAPQKGRPRFGKANKQAE 180
Db 144 LLCFGAAPSRSKRVLRVASCCKRLTRFHNQSELKDFGPTETAPQKGRPRFGKANKQAE 203
Qy 181 LENAY 185
Db 204 LENAY 208

RESULT 4
US-09-867-274-2
; Sequence 2, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; APPLICANT: Gao, Yongming
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-867-274-2

Query Match      17.3%; Score 32; DB 10; Length 139;
Best Local Similarity 100.0%; Pred. No. 1.6e-24;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 65 TDGFCRSKPVTELVCSGCGPARLLPNAIGR 96
Db 16 TDGFCRSKPVTELVCSGCGPARLLPNAIGR 47

RESULT 5
US-09-867-274-2
; Sequence 2, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; APPLICANT: Gao, Yongming
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-867-274-2

Query Match      17.3%; Score 32; DB 10; Length 190;
Best Local Similarity 100.0%; Pred. No. 2.1e-24;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

us-09-867-274-4.oligo.rapb

Mon Mar 31 09:44:33 2003

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:22:02 ; Search time 11.84 seconds

(without alignments)
917.557 Million cell updates/sec

Title: US-09-867-274-4
Perfect score: 185
Sequence: 1 QGWAQFRNDATVIGLGEY.....KGRKPRFGAKANQAELENAY 185

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 237916 seqs, 58723674 residues

Word size : 0

Total number of hits satisfying chosen parameters: 237916

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB_PEP.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB_PEP.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB_PEP.*
- 4: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB_PEP.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB_PEP.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCT05_PUBCOMB_PEP.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB_PEP.*
- 8: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB_PEP.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB_PEP.*
- 10: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB_PEP.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB_PEP.*
- 12: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB_PEP.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB_PEP.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB_PEP.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	185	100.0	185	US-09-867-274-4	Sequence 4, Appli
2	185	100.0	208	US-09-867-274-6	Sequence 6, Appli
3	32	17.3	139	US-09-864-761-47109	Sequence 47109, A
4	32	17.3	190	US-09-867-274-2	Sequence 2, Appli
5	32	17.3	213	US-09-867-274-5	Sequence 5, Appli
6	7	3.8	441	US-09-867-274-2	Sequence 2, Appli
7	7	3.8	535	US-09-815-242-13801	Sequence 13801, A
8	6	3.2	26	US-09-864-761-44917	Sequence 44917, A
9	6	3.2	40	US-09-864-761-40983	Sequence 40983, A
10	6	3.2	41	US-09-864-761-34998	Sequence 34998, A
11	6	3.2	46	US-09-925-297-836	Sequence 836, App
12	6	3.2	53	US-10-001-857-164	Sequence 164, App
13	6	3.2	56	US-09-876-478-16	Sequence 16, Appl
14	6	3.2	59	US-09-510-332-72	Sequence 72, Appl
15	6	3.2	70	US-09-864-761-35511	Sequence 35511, A
16	6	3.2	74	US-09-925-297-878	Sequence 878, App
17	6	3.2	83	US-09-790-264-17	Sequence 17, Appl
18	6	3.2	91	US-09-864-761-47343	Sequence 47343, A
19	6	3.2	110	US-09-741-669-320	Sequence 320, App

20	6	3.2	110	US-09-912-020-387	Sequence 387, App
21	6	3.2	110	US-09-815-242-10354	Sequence 10354, A
22	6	3.2	110	US-09-815-242-13970	Sequence 13970, A
23	6	3.2	112	US-09-815-242-5235	Sequence 5235, Ap
24	6	3.2	113	US-10-013-379-17	Sequence 17, Appl
25	6	3.2	117	US-09-815-242-12258	Sequence 12258, A
26	6	3.2	117	US-10-036-869-25	Sequence 12765, A
27	6	3.2	125	US-09-867-550-832	Sequence 25, Appl
28	6	3.2	137	US-09-925-297-879	Sequence 832, App
29	6	3.2	138	US-09-764-870-521	Sequence 521, App
30	6	3.2	139	US-09-815-242-11669	Sequence 11669, A
31	6	3.2	146	US-09-764-877-1316	Sequence 1316, Ap
32	6	3.2	160	US-09-815-242-11664	Sequence 11664, A
33	6	3.2	163	US-10-002-769-8	Sequence 8, Appli
34	6	3.2	181	US-10-101-464A-587	Sequence 587, App
35	6	3.2	186	US-09-944-160-16	Sequence 16, Appl
36	6	3.2	191	US-09-764-870-417	Sequence 417, App
37	6	3.2	195	US-09-858-664A-11	Sequence 11, Appl
38	6	3.2	199	US-10-140-293-12	Sequence 12, Appl
39	6	3.2	203	US-09-815-242-5050	Sequence 5050, Ap
40	6	3.2	205	US-09-905-291A-23	Sequence 23, Appl
41	6	3.2	205	US-10-066-500-99	Sequence 99, Appl
42	6	3.2	205	US-09-902-853-23	Sequence 23, Appl
43	6	3.2	205	US-09-907-824-23	Sequence 23, Appl
44	6	3.2	205	US-09-907-841-23	Sequence 23, Appl
45	6	3.2	205	US-09-907-841-23	Sequence 23, Appl

ALIGNMENTS

RESULT 1
US-09-867-274-4
; Sequence 4, Application US/09867274
; Patent No. US20020106650A1

; GENERAL INFORMATION: Christopher
; APPLICANT: Gao, Yongming
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-867-274-4

Query Match	100.0%	Score 185;	DB 10;	Length 185;
Best Local Similarity	100.0%;	Pred. No. 1.8e-177;		
Matches 185;	Conservative	0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	QGWAFRNDATVIGLGEYPPPPNNQTNWRAENGRRPPHPHYDAKDVSSEYSCRELHY	60	
Db	1	QGWAFRNDATVIGLGEYPPPPNNQTNWRAENGRRPPHPHYDAKDVSSEYSCRELHY	60	
QY	61	TRFLTGPCRSAPVTVELVCSQCGPARLLPNAIGRVKWRPNPDPFCIPDRYARQVRVQ	120	
Db	61	TRFLTGPCRSAPVTVELVCSQCGPARLLPNAIGRVKWRPNPDPFCIPDRYARQVRVQ	120	
QY	121	LLCPGGAAPRSRRKRVLVASCKKRLTRFHNSQLKDFGPETARPOKGRKPRFGAKANQAE	180	
Db	121	LLCPGGAAPRSRRKRVLVASCKKRLTRFHNSQLKDFGPETARPOKGRKPRFGAKANQAE	180	
QY	181	LENAY 185		
Db	181	LENAY 185		

```

RESULT 15
US-08-600-993A-30
; Sequence 30, Application US/08600993A
; Patent No. 5981229
; GENERAL INFORMATION:
; APPLICANT: Masure, H Robert
; APPLICANT: Pearce, Barbara J
; APPLICANT: Tuomanen, Elaine
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
; TITLE OF INVENTION: ACELLULAR VACCINES BASED THEREON
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/600,993A
; FILING DATE: 1-MAR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/245,511
; FILING DATE: 18-MAY-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/116,541
; FILING DATE: 01-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-069 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 97 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Bacillus subtilis
; US-08-600-993A-30

```

```

Query Match      3.2%; Score 6; DB 2; Length 97;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 167 GRKRP 172
Db 80 GRKRP 85

```

Search completed: March 28, 2003, 14:24:27
 Job time : 11.3467 secs

us-09-867-274-4.oligo.ra1

Mon Mar 31 09:44:32 2003

STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-836-791-5
Query Match 3.2%; Score 6; DB 2; Length 68;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 152 SELKDF 157
DB 35 SELKDF 40

RESULT 14
US-08-245-511-30
; Sequence 30, Application US/08245511
; Patent No. 5928900
; GENERAL INFORMATION:
; APPLICANT: Masure, H Robert
; APPLICANT: Pearce, Barbara J
; APPLICANT: Tuomanen, Elaine
; TITLE OF INVENTION: BACTERIAL EXPORTED PROTEINS AND
; TITLE OF INVENTION: ACCELLULAR VACCINES BASED THEREON
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/245,511
; FILING DATE: 18-MAY-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/116,541
; FILING DATE: 01-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-069 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 97 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Bacillus subtilis
US-08-245-511-30

Query Match 3.2%; Score 6; DB 2; Length 97;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 167 GRKPRP 172
DB 80 GRKPRP 85

NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: HU91-06A2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 24 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-220-272A-6

Query Match 3.2%; Score 6; DB 1; Length 24;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 AAPRSR 132
DB 13 AAPRSR 18
RESULT 13
US-08-836-791-5
; Sequence 5, Application US/08836791
; Patent No. 5886150
; GENERAL INFORMATION:
; APPLICANT: Duchesne, Marc
; APPLICANT: Faucher, Didier
; APPLICANT: Parker, Fabienne
; APPLICANT: Schweighoffer, Fabien
; APPLICANT: Tocque, Bruno
; TITLE OF INVENTION: Peptides Capable of Binding to the GAP
; TITLE OF INVENTION: Protein SH3 Domain, Nucleotide Sequences Coding Therefor,
; TITLE OF INVENTION: and Preparation and Use Thereof
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, Mailstop 3c43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,791
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 94/13955
; FILING DATE: 22-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 95/05753
; FILING DATE: 16-MAY-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO FR95/01539
; FILING DATE: 22-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Fehner Esq., Paul F.
; REGISTRATION NUMBER: 35,135
; REFERENCE/DOCKET NUMBER: ST94083G1-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 68 amino acids
; TYPE: amino acid

QY 127 AAPRSR 132
DB 13 AAPRSR 18
RESULT 13
US-08-836-791-5
; Sequence 5, Application US/08836791
; Patent No. 5886150
; GENERAL INFORMATION:
; APPLICANT: Duchesne, Marc
; APPLICANT: Faucher, Didier
; APPLICANT: Parker, Fabienne
; APPLICANT: Schweighoffer, Fabien
; APPLICANT: Tocque, Bruno
; TITLE OF INVENTION: Peptides Capable of Binding to the GAP
; TITLE OF INVENTION: Protein SH3 Domain, Nucleotide Sequences Coding Therefor,
; TITLE OF INVENTION: and Preparation and Use Thereof
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Inc.
; STREET: 500 Arcola Road, Mailstop 3c43
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/836,791
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 94/13955
; FILING DATE: 22-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 95/05753
; FILING DATE: 16-MAY-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO FR95/01539
; FILING DATE: 22-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Fehner Esq., Paul F.
; REGISTRATION NUMBER: 35,135
; REFERENCE/DOCKET NUMBER: ST94083G1-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (610) 454-3839
; TELEFAX: (610) 454-3808
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 68 amino acids
; TYPE: amino acid

QY 167 GRKPRP 172
DB 80 GRKPRP 85

Query Match 3.2%; Score 6; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 PGLGEY 20
Db 4 PGLGEY 9

RESULT 10

US-08-250-789A-96
; Sequence 96, Application US/08250789A
; Patent No. 5635597
; GENERAL INFORMATION:
; APPLICANT: Barrett, Ronald W.
; APPLICANT: Chernov-Rogan, Tania
; APPLICANT: Davis, Ann M.
; TITLE OF INVENTION: Peptides That Bind to IL-2 Receptors
; NUMBER OF SEQUENCES: 194
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/250,789A
; FILING DATE: 27-MAY-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5635597v1el, Vernon A.
; REGISTRATION NUMBER: 32,483
; REFERENCE/DOCKET NUMBER: 16528A-57/1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 96:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-250-789A-96

Query Match 3.2%; Score 6; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 PGLGEY 20
Db 4 PGLGEY 9

RESULT 11

US-08-250-789A-97
; Sequence 97, Application US/08250789A
; Patent No. 5635597
; GENERAL INFORMATION:
; APPLICANT: Barrett, Ronald W.
; APPLICANT: Chernov-Rogan, Tania
; APPLICANT: Davis, Ann M.
; TITLE OF INVENTION: Peptides That Bind to IL-2 Receptors
; NUMBER OF SEQUENCES: 194
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew

; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/250,789A
; FILING DATE: 27-MAY-1994
; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5635597v1el, Vernon A.
; REGISTRATION NUMBER: 32,483
; REFERENCE/DOCKET NUMBER: 16528A-57/1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 97:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-250-789A-97

Query Match 3.2%; Score 6; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 PGLGEY 20
Db 4 PGLGEY 9

RESULT 12

US-08-220-272A-6
; Sequence 6, Application US/08220272A
; Patent No. 5783384
; GENERAL INFORMATION:
; APPLICANT: Verdine, Gregory L.
; TITLE OF INVENTION: Selection of Binding Molecule
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02173

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/220,272A
; FILING DATE: 30-MAR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/029,331
; FILING DATE: February 10, 1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/819,855
; FILING DATE: January 13, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/00321
; FILING DATE: January 13, 1993
; ATTORNEY/AGENT INFORMATION:

us-09-867-274-4.oligo.ra1

Mon Mar 31 09:44:32 2003

FILE REFERENCE: 240083.508
 CURRENT APPLICATION NUMBER: US/09/449.218D
 CURRENT FILING DATE: 1999-11-24
 NUMBER OF SEQ ID NOS: 45
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 8
 LENGTH: 213
 TYPE: PRT
 ORGANISM: Homo sapien
 US-09-449-218D-8

Query Match 17.3%; Score 32; DB 4; Length 213;
 Best Local Similarity 100.0%; Pred. No. 3.8e-24;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 TDGFCRSKPVTELVCSCGCGPARLLPNAIGR 96
 Db 90 TDGFCRSKPVTELVCSCGCGPARLLPNAIGR 121

RESULT 7

US-09-449-218D-10

Sequence 10, Application US/09449218D

Patent No. 6395511

GENERAL INFORMATION:

APPLICANT: Brunkow, Mary E.

APPLICANT: Galas, David J.

APPLICANT: Kovacevich, Brian

APPLICANT: Mulligan, John T.

APPLICANT: Paepker, Bryan W.

APPLICANT: Van Ness, Jeffrey

APPLICANT: Winkler, David G.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING

TITLE OF INVENTION: BONE MINERALIZATION

FILE REFERENCE: 240083.508

CURRENT APPLICATION NUMBER: US/09/449.218D

CURRENT FILING DATE: 1999-11-24

NUMBER OF SEQ ID NOS: 45

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 10

LENGTH: 213

TYPE: PRT

ORGANISM: Cercopithecus pygerythrus

US-09-449-218D-10

Query Match 17.3%; Score 32; DB 4; Length 213;
 Best Local Similarity 100.0%; Pred. No. 3.8e-24;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 TDGFCRSKPVTELVCSCGCGPARLLPNAIGR 96
 Db 90 TDGFCRSKPVTELVCSCGCGPARLLPNAIGR 121

RESULT 8

US-08-250-789A-94

Sequence 94, Application US/08250789A

Patent No. 5635597

GENERAL INFORMATION:

APPLICANT: Barrett, Ronald W.

APPLICANT: Chernov-Rogan, Tania

APPLICANT: Davis, Ann M.

TITLE OF INVENTION: Peptides That Bind to IL-2 Receptors

NUMBER OF SEQUENCES: 194

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend Kourie and Crew

STREET: One Market Plaza, Steuart Tower, Suite 2000

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94105

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/250.789A
 FILING DATE: 27-MAY-1994
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 5635597viel, Vernon A.
 REGISTRATION NUMBER: 32,483
 REFERENCE/DOCKET NUMBER: 16528A-57/1043
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-326-2400
 TELEFAX: 415-326-2422
 INFORMATION FOR SEQ ID NO: 94:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 10 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-250-789A-94

Query Match 3.2%; Score 6; DB 1; Length 10;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 PGLGEY 20
 Db 4 PGLGEY 9

US-08-250-789A-95
 Sequence 95, Application US/08250789A
 Patent No. 5635597
 GENERAL INFORMATION:
 APPLICANT: Barrett, Ronald W.
 APPLICANT: Chernov-Rogan, Tania
 APPLICANT: Davis, Ann M.
 TITLE OF INVENTION: Peptides That Bind to IL-2 Receptors
 NUMBER OF SEQUENCES: 194
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Townsend and Townsend Kourie and Crew
 STREET: One Market Plaza, Steuart Tower, Suite 2000
 CITY: San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94105
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/250.789A
 FILING DATE: 27-MAY-1994
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 5635597viel, Vernon A.
 REGISTRATION NUMBER: 32,483
 REFERENCE/DOCKET NUMBER: 16528A-57/1043
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-326-2400
 TELEFAX: 415-326-2422
 INFORMATION FOR SEQ ID NO: 95:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 10 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-250-789A-95

RESULT 9
 US-08-250-789A-95
 Sequence 95, Application US/08250789A
 Patent No. 5635597
 GENERAL INFORMATION:
 APPLICANT: Barrett, Ronald W.
 APPLICANT: Chernov-Rogan, Tania
 APPLICANT: Davis, Ann M.
 TITLE OF INVENTION: Peptides That Bind to IL-2 Receptors
 NUMBER OF SEQUENCES: 194
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Townsend and Townsend Kourie and Crew
 STREET: One Market Plaza, Steuart Tower, Suite 2000
 CITY: San Francisco
 STATE: California
 COUNTRY: USA
 ZIP: 94105
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/250.789A
 FILING DATE: 27-MAY-1994
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 5635597viel, Vernon A.
 REGISTRATION NUMBER: 32,483
 REFERENCE/DOCKET NUMBER: 16528A-57/1043
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-326-2400
 TELEFAX: 415-326-2422
 INFORMATION FOR SEQ ID NO: 95:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 10 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-250-789A-95

; APPLICANT: Mulligan, John T.
; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-449-218D-14

Query Match 58.4%; Score 108; DB 4; Length 213;
Best Local Similarity 100.0%; Pred. No. 1.4e-99;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 65 TDGCRSAKPVTELVCSGCGPARLLPNAIGRVKWRPNPGRFCIPDRYRAQVOLLCP 124
Db 90 TDGCRSAKPVTELVCSGCGPARLLPNAIGRVKWRPNPGRFCIPDRYRAQVOLLCP 149
Qy 125 GGAAPRSRKVLVASCKKRLTRFNQSELKDFGPETARPQGRKPRP 172
Db 150 GGAAPRSRKVLVASCKKRLTRFNQSELKDFGPETARPQGRKPRP 197

RESULT 3
US-09-449-218D-16
; Sequence 16, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 176
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-449-218D-16

Query Match 17.3%; Score 32; DB 4; Length 176;
Best Local Similarity 100.0%; Pred. No. 3.2e-24;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 65 TDGCRSAKPVTELVCSGCGPARLLPNAIGR 96
Db 59 TDGCRSAKPVTELVCSGCGPARLLPNAIGR 90

RESULT 4
US-09-449-218D-2
; Sequence 2, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.

; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-449-218D-2

Query Match 17.3%; Score 32; DB 4; Length 213;
Best Local Similarity 100.0%; Pred. No. 3.8e-24;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 65 TDGCRSAKPVTELVCSGCGPARLLPNAIGR 96
Db 90 TDGCRSAKPVTELVCSGCGPARLLPNAIGR 121

RESULT 5
US-09-449-218D-6
; Sequence 6, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-449-218D-6

Query Match 17.3%; Score 32; DB 4; Length 213;
Best Local Similarity 100.0%; Pred. No. 3.8e-24;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 65 TDGCRSAKPVTELVCSGCGPARLLPNAIGR 96
Db 90 TDGCRSAKPVTELVCSGCGPARLLPNAIGR 121

RESULT 6
US-09-449-218D-8
; Sequence 8, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-449-218D-8

us-09-867-274-4.oligo.ra1

Mon Mar 31 09:44:32 2003

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:20:52 ; Search time 11.3467 Seconds
(without alignments)
479.722 Million cell updates/sec

Title: US-09-867-274-4
Perfect score: 185
Sequence: 1 CQWQAFRNDATVIGLGEY.....KGRKPRGAKANQALENAY 185

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size : 0

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/1/iaa/5A COMB.pcp.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pcp.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pcp.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pcp.*
5: /cgn2_6/ptodata/1/iaa/PCFUS COMB.pcp.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	174	94.1	211	4	US-09-449-218D-12
2	108	58.4	213	4	US-09-449-218D-14
3	32	17.3	176	4	US-09-449-218D-16
4	32	17.3	213	4	US-09-449-218D-2
5	32	17.3	213	4	US-09-449-218D-6
6	32	17.3	213	4	US-09-449-218D-8
7	32	17.3	213	4	US-09-449-218D-10
8	6	3.2	10	1	US-08-250-789A-94
9	6	3.2	10	1	US-08-250-789A-95
10	6	3.2	10	1	US-08-250-789A-96
11	6	3.2	10	1	US-08-250-789A-97
12	6	3.2	24	1	US-08-220-272A-6
13	6	3.2	68	2	US-08-836-791-5
14	6	3.2	97	2	US-08-245-511-30
15	6	3.2	97	2	US-08-600-993A-30
16	6	3.2	123	4	US-09-562-737-126
17	6	3.2	123	4	US-09-134-001C-3283
18	6	3.2	125	3	US-08-866-798-2
19	6	3.2	128	1	US-08-892-652-2
20	6	3.2	128	1	US-09-096-071-2
21	6	3.2	129	4	US-09-325-932A-50
22	6	3.2	182	4	US-09-230-670C-6
23	6	3.2	186	4	US-09-475-316A-13
24	6	3.2	187	4	US-09-134-001C-2866
25	6	3.2	199	3	US-08-737-248-7
26	6	3.2	199	3	US-08-325-932A-46
27	27	6	3.2	4	

Sequence 127, Appl
Sequence 13, Appl
Sequence 13, Appl
Sequence 3806, Appl
Sequence 3, Appl
Sequence 10, Appl
Sequence 10, Appl
Sequence 27, Appl
Sequence 3104, Appl
Sequence 33, Appl
Sequence 2, Appl
Sequence 94, Appl
Sequence 144, Appl
Sequence 278, Appl
Sequence 13, Appl
Sequence 13, Appl
Sequence 1, Appl
Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-449-218D-12
; Sequence 12, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John I.
; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; TITLE OF INVENTION: BONE MINERALIZATION
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 211
; TYPE: EXT
; ORGANISM: Mus musculus
US-09-449-218D-12

Query Match 94.1%; Score 174; DB 4; Length 211;
Best Local Similarity 100.0%; Pred. No. 4.5e-165; Indels 0; Gaps 0;
Matches 174; Conservative 0; Mismatches 0;
QY 1 CQWQAFRNDATVIGLGEYPPPPENNNTMRAENGRRPPHPHYDAKOVSEYSCRELHY 60
DB 24 CQWQAFRNDATVIGLGEYPPPPENNNTMRAENGRRPPHPHYDAKOVSEYSCRELHY 83
QY 61 TRFLTDGCRSAKPYTELVCSCQCGPARLLPNAIGRVKWRNENGPFCIPDRYRAQRVQ 120
DB 84 TRFLTDGCRSAKPYTELVCSCQCGPARLLPNAIGRVKWRNENGPFCIPDRYRAQRVQ 143
QY 121 LLCFGAAPSRRKRVLRVASCCKLRTFRHNSKLDKFGPETARPKGRKPRPGA 174
DB 144 LLCFGAAPSRRKRVLRVASCCKLRTFRHNSKLDKFGPETARPKGRKPRPGA 197

RESULT 2
US-09-449-218D-14
; Sequence 14, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Marin E., Nussbaum L., Vavasseur A., Forestier C.;
 RT "POP: an Arabidopsis NBD-like encoding sequence.";
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RX MEDLINE=97471969; PubMed=9330910;
 RA Sato S., Kotani H., Nakamura Y., Kaneko T., Asamizu E., Fukami M.,
 RA Miyajima N., Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence
 RT features of the 1.6 Mb regions covered by twenty physically assigned
 RT Pl clones.";
 RL DNA Res. 4:215-230(1997).
 DR EMBL; AF127664; AAD20643.1; -;
 DR EMBL; AB005239; BAB10978.1; -;
 DR InterPro; IPR003439; ABC_transportr.
 DR Pfam; PF00005; ABC_cran; 1.
 SQ SEQUENCE 282 AA; 31360 MW; 6F0B88F3EFDB1577 CRC64;

Query Match 3.8%; Score 7; DB 10; Length 282;
 Best Local Similarity 100.0%; Pred. No. 54;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 76 TELVCSG 82
 Db 82 TELVCSG 88

Search completed: March 28, 2003, 14:23:24
 Job time : 28.1467 secs

Mon Mar 31 09:44:34 2003

us-09-867-274-4.oligo.rspt

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 VIPGLGE 19
 ID |||||
 Db 222 VIPGLGE 228

RESULT 13

Q9LINO PRELIMINARY; PRT; 251 AA.
 AC Q9LINO; 2000 (TRENBLrel. 15, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE Putative DNA methylase.
 GN SCO6885 OR SC7F9.37.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Seeger K.J., Harris D.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

[2]

RP SEQUENCE FROM N.A.

RC STRAIN=A3(2);

RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;

RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=A3(2);

RX MEDLINE=97000351; PubMed=8843436;

RA Redenbach M., Kieser H.M., Denapante D., Eichner A., Cullum J.,

RA Kinashi H., Hopwood D.A.;

RT "A set of ordered cosmids and a detailed genetic and physical map for

the 8 Mb Streptomyces coelicolor A3(2) chromosome.";

RL Mol. Microbiol. 21:77-96(1996).

RN [4]

RP SEQUENCE FROM N.A.

RC STRAIN=A3(2) / M145;

RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,

RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,

RA Harper D., Batenan A., Brown S., Chandra G., Chen C.W., Collins M.,

RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,

RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,

RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,

RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,

RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,

RA Hopwood D.A.;

RT "Complete genome sequence of the model actinomycete Streptomyces

coelicolor A3(2).";

RL Nature 417:141-147(2002).

DR EMBL; AL138977; CAB2330.1; --

DR InterPro; IPR001091; C44 Mettransf.

DR InterPro; IPR002295; D21N6_mtfase.

DR InterPro; IPR002941; N6/N4_Mtase.

DR InterPro; IPR002052; N6 Mtase.

DR Pfam; PF01555; N6 N4 Mtase; 1.

DR PRINTS; PR00506; D21N6MTFRASE.

DR PRINTS; PR00508; S21N4MTFRASE.

DR PROSITE; PS00092; N6_MTASE; UNKNOWN_1.

KW Methyltransferase.

SQ SEQUENCE 251 AA; 27405 MW; 1AEC919125AF9CE9 CRC64;

Query Match 3.8%; Score 7; DB 16; Length 251;

Best Local Similarity 100.0%; Pred. No. 49;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 162 ARPOKGR 168

Db 127 ARPOKGR 133

RESULT 14

Q9KZ52 PRELIMINARY; PRT; 278 AA.
 ID Q9KZ52

AC Q9KZ52;

DT 01-OCT-2000 (TRENBLrel. 15, Created)

DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)

DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)

DE Putative insertion element transposase.

GN SCO7080 OR SC3A4.06C.

OS Streptomyces coelicolor.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.

OX NCBI_TaxID=1902;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=A3(2);

RA Brown S.P., Harris D.;

RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=A3(2);

RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;

RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=A3(2);

RX MEDLINE=97000351; PubMed=8843436;

RA Redenbach M., Kieser H.M., Denapante D., Eichner A., Cullum J.,

RA Kinashi H., Hopwood D.A.;

RT "A set of ordered cosmids and a detailed genetic and physical map for

the 8 Mb Streptomyces coelicolor A3(2) chromosome.";

RL Mol. Microbiol. 21:77-96(1996).

RN [4]

RP SEQUENCE FROM N.A.

RC STRAIN=A3(2) / M145;

RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,

RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,

RA Harper D., Batenan A., Brown S., Chandra G., Chen C.W., Collins M.,

RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,

RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,

RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,

RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,

RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,

RA Hopwood D.A.;

RT "Complete genome sequence of the model actinomycete Streptomyces

coelicolor A3(2).";

RL Nature 417:141-147(2002).

DR EMBL; AL138977; CAB2330.1; --

DR InterPro; IPR002295; D21N6_mtfase.

DR InterPro; IPR002941; N6/N4_Mtase.

DR InterPro; IPR002052; N6 Mtase.

DR Pfam; PF01555; N6 N4 Mtase; 1.

DR PRINTS; PR00506; D21N6MTFRASE.

DR PRINTS; PR00508; S21N4MTFRASE.

DR PROSITE; PS00092; N6_MTASE; UNKNOWN_1.

KW Methyltransferase.

SQ SEQUENCE 278 AA; 30918 MW; 81564528BF1A2E12 CRC64;

Query Match 3.8%; Score 7; DB 16; Length 278;

Best Local Similarity 100.0%; Pred. No. 54;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 168 RKPRPGA 174

Db 209 RKPRPGA 215

RESULT 15

Q9XF19 PRELIMINARY; PRT; 282 AA.
 ID Q9XF19

AC Q9XF19;

DT 01-NOV-1999 (TRENBLrel. 12, Created)

DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)

DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)

DE NBD-like protein.

GN POP.

OS Arabidopsis thaliana (Mouse-ear cross).

GN NM1656.
 OS Neisseria meningitidis (serogroup B).
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=491;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=MC58 / SEROGROUP B;
 RX MEDLINE=20175755; PubMed=10710307;
 RA Tetelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
 RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
 RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
 RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
 RA Mason T., Ciecko A., Parksey D.S., Blair E., Ciftone H., Clark E.B.,
 RA Corton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
 RA Gill J., Scarlato V., Masiagnani V., Pizzo M., Grandi G., Sun L.,
 RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
 RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
 MC58.";
 RL Science 287:1809-1815(2000).
 DR EMBL; AE002516; AAF42005.1; -.
 DR TIGR; NM1656; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 197 AA; 21804 MW; B6C59F55DEAESAC8 CRC64;

Query Match 3.8%; Score 7; DB 16; Length 197;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 175 KANQAE181
 DB 122 KANQAE123

RESULT 10

Q9JTA0
 ID Q9JTA0 PRELIMINARY; PRT; 197 AA.
 AC Q9JTA0
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE Hypothetical protein NMA1913.
 GN NMA1913.
 OS Neisseria meningitidis (serogroup A).
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=65699;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=Z2491 / SEROGROUP A / SEROTYPE 4A;
 RX MEDLINE=20222556; PubMed=10761919;
 RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
 RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
 RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
 RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
 RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
 RA Whitehead S., Spratt B.G., Barrall B.G.;
 RT "Complete DNA sequence of a serogroup A strain of Neisseria
 meningitidis Z2491.";
 RL Nature 404:502-506(2000).
 DR EMBL; AL162757; CAB85134.1; -.
 DR KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 197 AA; 21747 MW; B6C5977D4CE35AC8 CRC64;

Query Match 3.8%; Score 7; DB 16; Length 197;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 175 KANQAE181
 DB 122 KANQAE128

RESULT 11

Q9IBJ1
 ID Q9IBJ1 PRELIMINARY; PRT; 239 AA.
 AC Q9IBJ1;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE R-LORF1.
 GN R-LORF1.
 OS Turkey herpesvirus.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Marek's disease-like viruses.
 OX NCBI_TaxID=10390;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP STRAIN=GA;
 RX MEDLINE=92237304; PubMed=1315048;
 RA Jones D., Lee L., Liu J.L., Kung H.J., Tillotson J.K.;
 RT "Marek disease virus encodes a basic-leucine zipper gene resembling
 the fos/jun oncogenes that is highly expressed in lymphoblastoid
 tumors.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:4042-4046(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GA;
 RA Lee L.F., Wu P., Sui D., Ren D., Kung H.J.;
 RT "The Complete UL Sequence of Serotype I Marek's Disease Virus.";
 RL Proc. Natl. Acad. Sci. U.S.A. 0:0-0(2000).
 DR EMBL; AF147806; AAF67204.1; -.
 DR EMBL; AF147806; AAF6792.1; -.
 DR InterPro: IPR000564; 2Fe2S; ferredoxin.
 DR PROSITE: PS00197; 2Fe2S-FERREDOXIN; UNKNOWN 1.
 SQ SEQUENCE 239 AA; 24655 MW; 922B2C2B2950F01D CRC64;

Query Match 3.8%; Score 7; DB 12; Length 239;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 CPGGAAP 129
 DB 139 CPGGAAP 145

RESULT 12

Q8RXX1
 ID Q8RXX1 PRELIMINARY; PRT; 241 AA.
 AC Q8RXX1;
 DT 01-JUN-2002 (TReMBLrel. 21, Created)
 DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE Putative uridine kinase (Fragment).
 GN AT5G40870.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RC SEQUENCE FROM N.A.
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
 RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
 RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 RA Miranda M., Narusaka N., Nguyen M., Palm C.J., Sakurai T., Satou M.,
 RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
 RA Theologis A.;
 RT "Arabidopsis Full Length cDNA Clones";
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY080631; AAL85977.1; -.
 DR KW Kinase.
 FT NON TER
 SQ SEQUENCE 241 AA; 27003 MW; 266CF01112CEBED5 CRC64;

Query Match 3.8%; Score 7; DB 10; Length 241;
 Best Local Similarity 100.0%; Pred. No. 48;

Mon Mar 31 09:44:34 2003

|||||
Db 6 PGAKANQ 12

RESULT 6

Q9HVX4 ID Q9HVX4 PRELIMINARY; PRT; 148 AA.
AC Q9HVX4
DT 01-MAR-2001 (T-EMBLrel. 16, Created)
DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
DT 01-OCT-2001 (T-EMBLrel. 18, Last annotation update)
DE Hypothetical protein PA4441.
GN PA4441.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OC NCBI_TaxID=287;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=ATCC 15692 / PA01;
RC MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Huynh W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RL Nature 406:959-964 (2000).
DR EMBL; AE004858; AAG07829.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 148 AA; 16418 MW; 4DD9B4B217EDD2B8 CRC64;

Query Match 3.8%; Score 7; DB 16; Length 148;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 ARLLPNA 93
|||||

Db 26 ARLLPNA 32

RESULT 7

Q8Z095 ID Q8Z095 PRELIMINARY; PRT; 162 AA.
AC Q8Z095
DT 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (T-EMBLrel. 20, Last annotation update)
DE Hypothetical protein Alr0203.
GN Alr0203.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OC NCBI_TaxID=103690;
OX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=21595285; PubMed=1175840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriiguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsunoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120."
RL DNA Res. 8:205-213 (2001).
DR EMBL; AP003581; BAB7727.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 162 AA; 18240 MW; 7B56B0F5E663C79 CRC64;

Query Match 3.8%; Score 7; DB 16; Length 162;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 PGAKANQ 178

RESULT 8

Q9AK55 ID Q9AK55 PRELIMINARY; PRT; 169 AA.
AC Q9AK55
DT 01-JUN-2001 (T-EMBLrel. 17, Created)
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE Putative transposase.
GN SC04062 OR 28CD60.28.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OC NCBI_TaxID=1902;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN=A3(2);
RC Harris D.;
RA Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RL [2]
RN SEQUENCE FROM N.A.
RP STRAIN=A3(2);
RC Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RA Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RL [3]
RN SEQUENCE FROM N.A.
RP STRAIN=A3(2);
RC MEDLINE=97000351; PubMed=8943436;
RA Redenbach W., Kieser H.M., Denapante D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome."
RL Mol. Microbiol. 21:77-96 (1996).
RN [4]
RN SEQUENCE FROM N.A.
RP STRAIN=A3(2) / M145;
RC Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S., Taylor K.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2)."
RL Nature 417:141-147 (2002).
DR EMBL; AL583944; CAC32333.1; -
DR InterPro; IPR002559; Transposase 11.
DR Pfam; PF01609; Transposase 11; 1-
SQ SEQUENCE 169 AA; 18628 MW; A69871DB6E1902FE CRC64;

Query Match 3.8%; Score 7; DB 16; Length 169;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 39 RPPHPY 45
|||||

Db 153 RPPHPY 159

RESULT 9

Q9JYB9 ID Q9JYB9 PRELIMINARY; PRT; 197 AA.
AC Q9JYB9
DT 01-OCT-2000 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (T-EMBLrel. 19, Last annotation update)
DE Hypothetical protein NMB1656.

RESULT 2
Q9PH55
ID Q9PH55 PRELIMINARY; PRT; 89 AA.
AC Q9PH55
DT 01-MAR-2001 (TREMELrel. 16, Created)
DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMELrel. 16, Last annotation update)
DE Genomic DNA, chromosome 5, TAC clone.K1120.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucots II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RT clones.";
RL DNA Res. 7:31-63 (2000).
DR EMBL; AB022211; BAB10713.1; -- DF12344C936FAEOA CRC64;
SQ SEQUENCE 89 AA; 10345 MW; 10345 MW; 10345 MW; 10345 MW; 10345 MW;

Query Match 3.8%; Score 7; DB 10; Length 89;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 88 RLLPNAI 94
DB 18 RLLPNAI 24

RESULT 3
Q8ZJS3
ID Q8ZJS3 PRELIMINARY; PRT; 89 AA.
AC Q8ZJS3
DT 01-MAR-2002 (TREMELrel. 20, Created)
DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMELrel. 20, Last annotation update)
DE Hypothetical protein YP00013A.
GN YP00013A.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / BIOVAR ORIENTALIS;
RX MEDLINE=21470413; PubMed=1158360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdono-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltham T., Hamlin N., Holtroyd S., Jagels K., Kariyeh A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527 (2001).
DR EMBL; AJ414141; CAC88881.1; --
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 89 AA; 10168 MW; C240C03A620D66F9 CRC64;

Query Match 3.8%; Score 7; DB 16; Length 89;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 11 TEVIPGL 17
DB 62 TEVIPGL 68

RESULT 4
Q9GTJ4
ID Q9GTJ4 PRELIMINARY; PRT; 109 AA.
AC Q9GTJ4
DT 01-MAR-2001 (TREMELrel. 16, Created)
DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMELrel. 21, Last annotation update)
DE Zinc metalloproteinase 3 MBP3 (Fragment).
OS Ancylostoma caninum (Dog hookworm).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC Ancylostomatidae; Ancylostomatinae; Ancylostominae; Ancylostoma.
OX NCBI_TaxID=29170;
RN [1]
RP SEQUENCE FROM N.A.
RA Jones B.F., Hotez P.J.;
RT "Cloning and characterization of a zinc-metalloprotease secreted by
RT the invasive stages of Ancylostoma caninum.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF273706; AAG23106.1; --
DR HSSP; P08473; 1DWT.
DR InterPro: IPR000718; Peptidase M13.
DR Pfam; PF01431; Peptidase M13; 1.
FT NON_TER 1
SQ SEQUENCE 109 AA; 12300 MW; 7C3C54FD724CAD7B CRC64;

Query Match 3.8%; Score 7; DB 5; Length 109;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 106 DFRCPID 112
DB 86 DFRCPID 92

RESULT 5
Q910CO
ID Q910CO PRELIMINARY; PRT; 130 AA.
AC Q910CO
DT 01-MAR-2001 (TREMELrel. 16, Created)
DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
DT 01-OCT-2001 (TREMELrel. 18, Last annotation update)
DE Hypothetical protein PA2722.
GN PA2722.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01,
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong K.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964 (2000).
DR EMBL; AE004700; AAG06110.1; --
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 130 AA; 14641 MW; ECEBDCD7F247836E CRC64;

Query Match 3.8%; Score 7; DB 16; Length 130;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 133 KVLVLAS 139
DB 12 KVLVLAS 18

Mon Mar 31 09:44:34 2003

GenCore version 5.1.4_p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:19:52 ; Search time 26.1467 Seconds
(without alignments)
1457.881 Million cell updates/sec

Title: US-09-867-274-4
Perfect score: 185
Sequence: 1 QGQAFRDATEVPLGLGVY.....KGRKPRGAKANQAELENAY 185

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size : 0 671580

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SPTREMBL 21.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	8	4.3	428	4 Q969X8	Q969X8 homo sapien
2	7	3.8	89	10 Q9FH55	Q9FH55 arabidopsis
3	7	3.8	89	16 Q8ZJS3	Q8ZJS3 yersinia pe
4	7	3.8	109	5 Q9GTJ4	Q9GTJ4 ancylostoma
5	7	3.8	130	16 Q9T0C0	Q9T0C0 pseudomonas
6	7	3.8	148	16 Q9HVX4	Q9HVX4 pseudomonas
7	7	3.8	162	16 Q8Z095	Q8Z095 anabaena sp
8	7	3.8	169	16 Q9AK55	Q9AK55 streptomyce
9	7	3.8	197	16 Q9JYB9	Q9JYB9 neisseria m
10	7	3.8	197	16 Q9JTA0	Q9JTA0 neisseria m
11	7	3.8	239	12 Q9IBJ1	Q9IBJ1 turkey heip
12	7	3.8	241	10 Q8RXX1	Q8RXX1 arabidopsis
13	7	3.8	251	16 Q9LJN0	Q9LJN0 streptomyce
14	7	3.8	278	16 Q9KZ52	Q9KZ52 streptomyce
15	7	3.8	282	10 Q9XF19	Q9XF19 arabidopsis
16	7	3.8	311	16 Q8UBG9	Q8UBG9 agrobacteri

Q9PQ72 ureaplasma
Q9FUN1 zea mays (m
Q945Y7 pseudomonas
Q91X3 streptomyce
Q91G89 oryza sativ
Q90388 methanococc
Q95K71 macaca fasc
Q19348 caenorhabdi
Q9VPQ8 drosophila
Q8S485 zea mays (m
Q91Y6 arabidopsis
Q91K34 arabidopsis
Q9F80 arabidopsis
Q9HVR9 pseudomonas
Q8ZLA9 salmonella
Q8Z299 salmonella
Q8X4K7 escherichia
Q9594 caenorhabdi
Q81837 arabidopsis
P95077 mycobacteri
Q44003 toxoplasma
Q92D38 listeria in
Q9H7Y7 homo sapien
Q85050 pseudorabie
Q9PGB8 xylella fas
Q9BSC4 homo sapien
Q8XPT2 raistonia s
Q9ABT1 caulobacter
Q90X95 xenopus lae

ALIGNMENTS

RESULT 1

Q969X8 PRELIMINARY; PRT; 428 AA.
ID Q969X8
AC Q969X8; 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Elki protein.
GN ELK1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=HIPPOCAMPUS;
RX MEDLINE=99246057; PubMed=10231026;
RA Yamauchi T., Toko M., Suga M., Hatakeyama T., Isobe M.;
RT "Structural organization of the human ELK1 gene and its processed
pseudogene ELK2 genes.";
RL DNA Res. 6:21-27(1999).
DR EMBL; AB016194; BAA36616.1; -;
DR EMBL; AB016193; BAA36616.1; -;
DR InterPro; IPR000418; Ets.
DR InterPro; IPR002341; HSP_ETs.
DR InterPro; IPR00178; Ets; 1.
DR Pfam; PF00178; ETS_DOMAIN_1; UNKNOWN 1.
DR PROSITE; PS00345; ETS_DOMAIN_2; UNKNOWN 1.
DR PROSITE; PS00346; ETS_DOMAIN_3; 1.
DR PROSITE; PS00061; ETS_DOMAIN_3; 1.
SQ SEQUENCE 428 AA; 47915 MW; 58FBEF8ADB985A83 CRC64;

Query Match 4.3%; Score 8; DB 4; Length 428;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 POKGRKPR 171

DB 310 POKGRKPR 317

Mon Mar 31 09:44:34 2003

CC SPASTIC PARALYSIS OF THE ANTERIOR AND POSTERIOR EXTREMITIES AND
 CC DEATH AT DOSE LEVELS OF 0.24 MG/MOUSE. INSECTICIDAL TO THE LARVAL
 CC AND ADULT FORMS OF THE HOUSE FLY.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Expressed by the venom gland.
 CC -!- SIMILARITY: BELONGS TO THE SPIDER TOXIN TX2 FAMILY.
 DR PIR; B39305; B39305.
 DR PIR; S29215; S29215.
 KW Sodium channel inhibitor; Toxin; Neurotoxin.
 SQ SEQUENCE 49 AA; 5111 MW; 77B46AAB3911716 CRC64;

Query Match 3.2%; Score 6; DB 1; Length 49;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 138 ASCKCK 143
 |
 |
 |
 |
 |
 Db 44 ASCKCK 49

Search completed: March 28, 2003, 14:22:22
 Job time : 8.89333 secs

CC OF MYOSIN I TO CROSS-LINK ACTIN FILAMENTS.
 CC -1- MISCELLANEOUS: THIS ORGANISM EXPRESSES AT LEAST THREE ISOFORMS OF
 CC MYOSIN I HEAVY-CHAIN, ENCODED BY GENES MIA, MIB, AND MIC.
 CC -1- SIMILARITY: BELONGS TO THE MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -1- CAUTION: WAS ORIGINALLY THOUGHT TO BE MYOSIN IB.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; J02974; AAA27707.1; -
 CC PIR; A33891; MWAXIC.
 CC HSP; P08799; 1MND.
 CC InterPro; IPR001452; SH3.
 CC InterPro; IPR001609; myosin_head.
 CC Pfam; PF00018; SH3; 1.
 CC Pfam; PF00063; myosin_head; 1.
 CC PRINTS; PR00193; MYOSINHEAVY.
 CC PRINTS; PR00452; SH3DOMAIN.
 CC ProDom; PD000066; myosin_head; 1.
 CC ProDom; PD000353; myosin_head; 1.
 CC SMART; SM00242; MYSC; 1.
 CC SMART; SM00326; SH3; 1.
 CC PROSITE; PS50002; SH3; 1.
 CC MYOSIN; ATP-binding; Phosphorylation; Multigene family; SH3 domain.
 CC DOMAIN 1 670
 CC MYOSIN HEAD-LIKE.
 CC DOMAIN 671 922
 CC TAIL HOMOLGY REGION 1 (TH.1).
 CC DOMAIN 923 975
 CC GLY/PRO/ALA-RICH (TH.2).
 CC DOMAIN 976 1035
 CC SH3.
 CC DOMAIN 1036 1168
 CC GLY/PRO/ALA-RICH (TH.2).
 CC NP_BIND 101 108
 CC ATP (POTENTIAL).
 CC MOD_RES 311 311
 CC PHOSPHORYLATION.
 CC SEQUENCE 1168 AA; 127309 MW; D07084B37A37A32 CRC64;
 CC
 CC Query Match 3.8%; Score 7; DB 1; Length 1168;
 CC Best Local Similarity 100.0%; Pred. No. 26;
 CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 124 PGGAAPR 130
 CC Db 1103 PGGAAPR 1109
 CC
 CC RESULT 14
 CC BCC3 ACXY
 CC ID_BCC3 ACXY STANDARD; PRT; 1325 AA.
 CC AC Q9WXE3;
 CC DT 15-JUN-2002 (Rel. 41, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last sequence update)
 CC DE Cellulose synthase 1 operon protein C precursor.
 CC BCC3.
 CC OS Acetobacter xylinus.
 CC AC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
 CC OC Gluconacetobacter.
 CC OX NCBI_TaxID=28448;
 CC [1]
 CC RN SEQUENCE FROM N.A.
 CC RP STRAIN=JCM 7664 / IFO 13693;
 CC RX PubMed=10382968;
 CC RA Umeda Y., Hirano A., Ishibashi M., Akiyama H., Onizuka T., Ikeuchi M.,
 CC Inoue Y.;
 CC RT "Cloning of cellulose synthase genes from Acetobacter xylinum JCM
 CC RT 7664: implication of a novel set of cellulose synthase genes.";
 CC RL DNA Res. 6:109-115(1999).
 CC -1- FUNCTION: Required for maximal bacterial cellulose synthesis. It
 CC may be involved in the formation of a membrane complex for
 CC extrusion of the cellulose product (By similarity).
 CC

CC -1- PATHWAY: Bacterial cellulose biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Outer membrane (Potential).
 CC -1- SIMILARITY: BELONGS TO THE ACSC/BCSC FAMILY.
 CC -1- SIMILARITY: CONTAINS 8 TPR REPEATS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AB015802; BAA77587.1; -
 CC InterPro; IPR003921; Cellsynth_C.
 CC InterPro; IPR001440; TPR.
 CC Pfam; PF00515; TPR; 4.
 CC PRINTS; PR01441; CELLSYNTHASEC.
 CC Cellulose biosynthesis; Repeat; TPR repeat; Membrane; Outer membrane;
 CC Signal.
 CC SIGNAL 1 30
 CC POTENTIAL.
 CC FT CHAIN 31 1325
 CC CELLULOSE SYNTHASE 1 OPERON PROTEIN C.
 CC FT REPEAT 50 93
 CC TPR 1.
 CC FT REPEAT 85 117
 CC TPR 2.
 CC FT REPEAT 292 325
 CC TPR 3.
 CC FT REPEAT 326 359
 CC TPR 4.
 CC FT REPEAT 406 439
 CC TPR 5.
 CC FT REPEAT 558 591
 CC TPR 6.
 CC FT REPEAT 702 735
 CC TPR 7.
 CC FT REPEAT 737 769
 CC TPR 8.
 CC SQ SEQUENCE 1325 AA; 142541 MW; DA4BEECFADDED74D CRC64;
 CC
 CC Query Match 3.8%; Score 7; DB 1; Length 1325;
 CC Best Local Similarity 100.0%; Pred. No. 29;
 CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 36 NGRGPPH 42
 CC Db 161 NGRGPPH 167
 CC
 CC RESULT 15
 CC TX25 PHONI
 CC ID_TX25 PHONI STANDARD; PRT; 49 AA.
 CC AC P29424;
 CC DT 01-APR-1993 (Rel. 25, Created)
 CC DT 01-APR-1993 (Rel. 25, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Neurotoxin Tx2-5.
 CC OS Phonetria nigriventer (Brazilian armed spider).
 CC OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
 CC OC Araneomorphae; Entelegynae; Lycosoidea; Ctenidae; Phonetria.
 CC OX NCBI_TaxID=6918;
 CC RN [1]
 CC RP SEQUENCE
 CC RC TISSUE=Venom;
 CC RX MEDLINE=93011905; PubMed=1397265;
 CC RA Cordeiro M.N., Diniz C.R., Valentim A.C., von Eickstedt V.R.D.,
 CC Gilroy J., Richardson M.;
 CC RT "The purification and amino acid sequences of four Tx2 neurotoxins
 CC RT from the venom of the Brazilian 'armed' spider Phonetria nigriventer
 CC RT (Keys).";
 CC RN FEBS Lett. 310:153-156(1992).
 CC [2]
 CC RP SEQUENCE OF 1-10.
 CC RC TISSUE=Venom;
 CC RX MEDLINE=92196803; PubMed=1801316;
 CC RA Rezende L. Jr., Cordeiro M.N., Oliveira E.B., Diniz C.R.;
 CC RT "Isolation of neurotoxic peptides from the venom of the 'armed'
 CC RT spider Phonetria nigriventer.";
 CC RL Toxicon 29:1225-1233(1991).
 CC -1- FUNCTION: BLOCKS VOLTAGE-GATED SODIUM CHANNELS. CAUSES SCRATCHING,
 CC LACRIMATION, HYPERSALIVATION, SWEATING AND AGITATION FOLLOWED BY

FT	CARBOHYD	398	.	N-LINKED (GLC.)	.
FT	CARBOHYD	398	.	N-LINKED (GLC.)	.
FT	CARBOHYD	438	.	N-LINKED (GLC.)	.
FT	CARBOHYD	513	.	N-LINKED (GLC.)	.
FT	CARBOHYD	643	.	N-LINKED (GLC.)	.

CC CC CC CC CC
MIOSINS I CAN INTERACT WITH F-ACTIN; THIS TOGETHER WITH
VERSAL TH.2 AND SH3 BIND TIGHTLY TO F-ACTIN; THIS TOGETHER WITH
THE NUCLEOTIDE-SENSITIVE SITE IN THE HEAD, ALLOWS SINGLE MOLECULES

FT	CARBOHYD	396	N-LINKED (GLC. . .)	396
FT	CARBOHYD	438	N-LINKED (GLC. . .)	438
FT	CARBOHYD	513	N-LINKED (GLC. . .)	513
FT	CARBOHYD	643	N-LINKED (GLC. . .)	643

RA Olson E.R., Donyak D.S., Jurss L.M., Poorman R.A.;
 RT "Identification and characterization of dppA, an *Escherichia coli*
 RL gene encoding a periplasmic dipeptide transport protein.";
 RN J. Bacteriol. 173:234-244(1991).
 RP [2]
 RC SEQUENCE FROM N.A., AND SEQUENCE OF 29-50.
 RX STRAIN-K12;
 RA MEDLINE=92065799; PubMed=1956284;
 RA Abouhanad W.N., Manson M., Gibson M.M., Higgins C.F.;
 RT "Peptide transport and chemotaxis in *Escherichia coli* and *Salmonella*
 RL typhimurium: characterization of the dipeptide permease (Dpp) and the
 RN dipeptide-binding protein";
 RP Mol. Microbiol. 5:1035-1047(1991).
 RA [3]
 RC SEQUENCE FROM N.A.
 RX STRAIN-K12 / MM500;
 RA MEDLINE=95231288; PubMed=7536291;
 RA Abouhanad W.N., Manson M.D.;
 RT "The dipeptide permease of *Escherichia coli* closely resembles other
 RL bacterial transport systems and shows growth-phase-dependent
 RN expression";
 RP Mol. Microbiol. 14:1077-1092(1994).
 RA [4]
 RC SEQUENCE FROM N.A.
 RX STRAIN-K12 / MG1655;
 RA MEDLINE=94316500; PubMed=8041620;
 RA Sofia H.J., Burland V., Daniels D.L., Plunkett G. III, Blattner F.R.;
 RT "Analysis of the *Escherichia coli* genome. V. DNA sequence of the
 RL region from 76.0 to 81.5 minutes.";
 RN Nucleic Acids Res. 22:2576-2586(1994).
 RA [5]
 RC SEQUENCE OF 29-40.
 RX STRAIN-K12 / EMG2;
 RA MEDLINE=97443975; PubMed=9298646;
 RA Link A.J., Robison K., Church G.M.;
 RT "Comparing the predicted and observed properties of proteins encoded
 RL in the genome of *Escherichia coli* K-12.";
 RN Electrophoresis 18:1259-1313(1997).
 RA [6]
 RC DISULFIDE BONDS.
 RX Dunten P.;
 RA Submitted (JUN-1995) to the SWISS-PROT data bank.
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (3.2 ANGSTROMS).
 RX MEDLINE=96155055; PubMed=8563629;
 RA Dunten P., Mowbray S.L.;
 RT "Crystal structure of the dipeptide binding protein from *Escherichia*
 RL coli involved in active transport and chemotaxis.";
 RN Protein Sci. 4:2327-2334(1995).
 RA [8]
 RC X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
 RX MEDLINE=96118375; PubMed=8527431;
 RA Nickitenko A.V., Trakhanov S., Quitocho F.A.;
 RT "2-A resolution structure of DppA, a periplasmic dipeptide
 RL transport/chemosensory receptor.";
 RN Biochemistry 34:16585-16595(1995).
 CC -!- FUNCTION: DIPEPTIDE-BINDING PROTEIN OF AN OSMOTIC-SHOCKABLE
 CC TRANSPORT SYSTEM. DPPA IS ALSO REQUIRED FOR PEPTIDE CHEMOTAXIS.
 CC -!- SUBCELLULAR LOCATION: Periplasmic.
 CC -!- SIMILARITY: BELONGS TO THE BACTERIAL EXTRACELLULAR SOLUTE-BINDING
 CC PROTEIN FAMILY 5.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M35045; AAA23707.1; -;
 CC EMBL; X58051; CAA41090.1; -;
 CC EMBL; L08399; AAA23702.1; -;

DR EMBL; U00039; AAB18522.1; -;
 DR EMBL; AE000431; AAC76569.1; -;
 DR PIR; S15292; S15292;
 DR PIR; A39194; A39194;
 DR PDB; 1DPE; 17-AUG-96.
 DR PDB; 1DPE; 17-AUG-96.
 DR SWISS-2DPAGE; P23847; COLI.
 DR ECO2DBASE; G059.9; 6TH EDITION.
 DR EcoGene; EG10248; dppA.
 DR InterPro; IPR000914; ssp_bac_5.
 DR Pfam; PF00496; ssp_bac_5_1_1.
 DR PROSITE; PS01040; SBP_BACTERIAL_5; 1.
 KW Peptide transport; Transport; Periplasmic; Signal; Chemotaxis;
 KW 3D-structure; Complete proteome.
 FT SIGNAL 1 38 PERIPLASMIC DIPEPTIDE TRANSPORT PROTEIN.
 FT CHAIN 29 535
 FT DISULFID 34 262
 FT DISULFID 450 463
 SQ SEQUENCE 535 AA; 60293 MW; 3F7C4756E8A8C2F0 CRC64;
 Query Match 3.8%; Score 7; DB 1; Length 535;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 11 TEVIPGL 17
 Db 72 TEVIPGL 78
 RESULT 11
 VGLG SYNV STANDARD; PRT; 632 AA.
 ID VGLG SYNV
 AC P27277;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 01-JUL-1993 (Rel. 26, Last annotation update)
 DE Spike glycoprotein precursor.
 GN G.
 OS Sonchus yellow net virus (SYNV).
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Rhabdoviridae; Nucleorhabdovirus.
 OX NCBI_TaxID=11307;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC PV-263;
 RX MEDLINE=92024089; PubMed=1926779;
 RA Goldberg K.B., Modrell B., Hillman B.I., Heaton L.A., Choi T.J.,
 RA Jackson A.O.;
 RT "Structure of the glycoprotein gene of sonchus yellow net virus, a
 RL plant rhabdovirus.";
 RL Virology 185:32-38(1991).
 CC -!- FUNCTION: THIS PROTEIN FORMS SPIKES ON THE SURFACE OF THE VIRION.
 CC IT IS RESPONSIBLE BOTH FOR THE BINDING OF THE VIRUS TO SUSCEPTIBLE
 CC HOST CELLS AND FOR INDUCING THE UPTAKE OF THE VIRUS BY THE CELL.
 CC THE INTERACTION BETWEEN THE INTERNAL COMPONENTS OF THE VIRION
 CC AND THE PORTION OF THE GLYCOPROTEIN EXPOSED ON THE CYTOPLASMIC
 CC FACE OF THE PLASMA MEMBRANE PROBABLY DIRECTS ENVELOPMENT AND
 CC VIRUS BUDDING.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L32603; AAA50384.1; -;
 CC EMBL; M73626; AAA47898.1; -;
 CC PIR; A40776; VGVNSY.
 KW Transmembrane; Envelope protein; Glycoprotein; Signal.
 FT SIGNAL 1 17 POTENTIAL.
 FT CHAIN 18 632 SPIKE GLYCOPROTEIN.

```
CC EMBL; X87257; CA660715.1; -
DR EMBL; Z36939; CA85391.1; -
DR HSSP; P28324; 1BC8.
DR TRANSFAC; T05013; -.
DR MGD; MGI:101833; Elk1.
DR InterPro; IPR000418; Ets.
DR InterPro; IPR002341; HSF_ETs.
DR Pfam; PF00178; Ets; 1.
DR PRINTS; PR00454; ETSDOMAIN.
DR SMART; SM00413; ETS; 1.
DR PROSITE; PS00345; ETS DOMAIN 1; 1.
DR PROSITE; PS00346; ETS DOMAIN 2; 1.
DR PROSITE; PS00661; ETS DOMAIN 3; 1.
DR Transcription regulation; Activator; Nuclear protein; DNA-binding;
KW Phosphorylation.
KW DNA_BIND 5 ETS-DOMAIN
FT CONFLICT 133 133 P -> T (IN REF. 2).
SQ SEQUENCE 429 AA; 45243 MW; B61B5B977731D54F CRC64;

Query Match 4.3%; Score 8; DB 1; Length 429;
Best Local Similarity 100.0%; Pred. No. 0.82;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 PQGKPKPR 171
DB 311 PQGKPKPR 318

RESULT 8
VE7 HPV44
ID VE7 HPV44 STANDARD; PRT; 97 AA.
AC Q80914;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE E7 protein.
GN E7.
OS Human papillomavirus type 44.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10592;
RN [1]
RP SEQUENCE FROM N.A.
RA Delius H.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: E7 PROTEIN HAS BOTH TRANSFORMING AND TRANS-ACTIVATING
CC ACTIVITIES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U31788; AAU79458.1; -.
DR EMBL; U31788; AAU79458.1; -.
DR InterPro; IPR000148; Papi_E7.
DR Pfam; PF00527; E7; 1.
DR Early protein; Transcription regulation; Oncogene;
KW DNA-binding; Trans-acting factor.
FT SITE 57 60 C-XX-C MOTIF-1.
FT SITE 90 93 C-XX-C MOTIF-2.
SQ SEQUENCE 97 AA; 10641 MW; B4866AE13B050456 CRC64;

Query Match 3.8%; Score 7; DB 1; Length 97;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 SRKRLV 137
DB 61 SRKRLV 67
```

```
RESULT 9
RL22 LEPIN
ID RL22 LEPIN STANDARD; PRT; 110 AA.
AC Q9XD31;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 50S ribosomal protein L22.
GN RPLV.
OS Leptospira interrogans.
OC Bacteria; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Serovar Lai;
RC MEDLINE=20088835; PubMed=10620683;
RA Zuercher R.L., Hartskeerl R.A., van de Kemp H., Bal A.E.;
RA "Characterization of the Leptospira interrogans S10-spc-alpha
operon.";
RT FEMS Microbiol. Lett. 182:303-308(2000).
CC -1- FUNCTION: THIS PROTEIN BINDS SPECIFICALLY TO 23S rRNA; ITS BINDING
CC IS STIMULATED BY OTHER RIBOSOMAL PROTEINS, E.G., L4, L17, AND L20.
CC IT IS IMPORTANT DURING THE EARLY STAGES OF 50S RECONSTITUTION
CC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE L22P FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF115283; AA040588.1; -.
DR HSSP; P48286; IBXE.
DR InterPro; IPR001063; Ribosomal_L22.
DR Pfam; PF00237; Ribosomal_L22; 1.
DR ProDom; PD001032; Ribosomal_L22; 1.
DR TIGRPFAMs; TIGR01044; rplV_bact; 1.
DR PROSITE; PS00464; RIBOSOMAL_L22; 1.
KW Ribosomal protein; rRNA-binding.
SQ SEQUENCE 110 AA; 12566 MW; B9181B85E9756CAA CRC64;

Query Match 3.8%; Score 7; DB 1; Length 110;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 RKRLVA 138
DB 15 RKRLVA 21

RESULT 10
DPPA ECOLI
ID DPPA ECOLI STANDARD; PRT; 535 AA.
AC P23847;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Periplasmic dipeptide transport protein precursor (Dipeptide-binding
DE protein) (DBP).
GN DPPA OR B3544.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RC MEDLINE=91100289; PubMed=1702779;
RX
```

```

DR HSP: P28324; lBC8.
DR TRANSFAC; T00250; -.
DR Genew; HGNC:3321; ELK1.
DR MIM; 311040; -.
DR InterPro; IPR000418; Ets.
DR InterPro; IPR002341; HSF_ETs.
DR Pfam; PF00176; Ets; 1.
DR PRINTS; PR00454; ETSDOMAIN.
DR SMART; SM00413; ETS; 1.
DR PROSITE; PS00345; ETS_DOMAIN_1; 1.
DR PROSITE; PS00346; ETS_DOMAIN_2; 1.
DR PROSITE; PS0061; ETS_DOMAIN_3; 1.
KW Transcription regulation; Activator; Nuclear protein; DNA-binding;
KW Phosphorylation; Alternative splicing.
FT DNA_BIND 5 86 ETS-DOMAIN.
FT VAREFLIC 91 95 SHCAP (IN ISOFORM 2).
FT VAREFLIC 96 428 MISSING (IN ISOFORM 2).
FT CONFLICT 183 193 S -> N (IN REF. 1).
SQ SEQUENCE 428 AA; 44888 MW; 68F71F8ADB9D38CA CRC64;

Query Match 4.3%; Score 8; DB 1; Length 428;
Best Local Similarity 100.0%; Pred. No. 0.82;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 164 PQGKPKR 171
    |||||
Db 310 PQGKPKR 317

RESULT 7
ELK1 MOUSE
ID ELK1 MOUSE STANDARD; PRT; 429 AA.
AC P41969;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE ETS-domain protein ELK-1.
GN ELK1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Embryo;
RA MEDLINE=97017146; PubMed=8863747;
RA Grevin D., Ung S., Denhez F., Dehem M., Quatannens B., Begue A.,
RA Stehelin D., Martin P.;
RT "Structure and organization of the mouse elk1 gene.";
RL Gene 174:195-188(1996).
RP [2]
RC SEQUENCE OF 5-224 FROM N.A.
RP TISSUE=Embryo;
RX MEDLINE=95047310; PubMed=7958835;
RA Giovane A., Pintzas A., Maira S.-M., Sobieszczuk P., Wasyluk B.;
RT "Net, a new ets transcription factor that is activated by Ras.";
RL Genes Dev. 8:1502-1513(1994).
CC -!- FUNCTION: STIMULATES TRANSCRIPTION. BINDS TO PURINE-RICH DNA
CC SEQUENCES. CAN FORM A TERNARY COMPLEX WITH THE SERUM RESPONSE
CC FACTOR AND THE ETS AND SRF MOTIFS OF THE POS SERUM RESPONSE
CC ELEMENT.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN THE BRAIN, AND TO A
CC LESSER EXTENT IN THE HEART, LIVER AND MUSCLE.
CC -!- SIMILARITY: BELONGS TO THE ETS FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

```

Mon Mar 31 09:44:34 2003

```

FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT NON_TER 176 176
SQ SEQUENCE 176 AA; 19743 MW; 35F0CA61A425F4DB CRC64;

Query Match 17.3%; Score 32; DB 1; Length 176;
Best Local Similarity 100.0%; Pred. No. 1.4e-27; Indels 0; Gaps 0;
Matches 32; Conservative 0; Mismatches 0;

Qy 65 TDGCRSAKPVTGLVCSGCGCPARLLPNAIGR 96
Db 59 TDGCRSAKPVTGLVCSGCGCPARLLPNAIGR 90

RESULT 4
SOST_CERAE STANDARD; PRT; 213 AA.
AC Q9BG78;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sclerostin precursor.
GN SOST.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21090529; PubMed=11179006;
RA Brunkow M.E., Gardner J.C., Van Ness J., Paepser B.W., Kovacevich B.R.,
RA Prohl S., Skonier J.E., Zhao L., Sabo P.J., Fu Y.H., Alisch R.S.,
RA Gillett L., Colbert T., Tacconi P., Galas D., Hamersma H.,
RA Beighton P., Mulligan J.T.;
RA "Bone dysplasia sclerosteosis results from loss of the SOST gene
RT product, a novel cysteine knot-containing protein.";
RL Am. J. Hum. Genet. 68:577-589(2001).
CC -!- FUNCTION: Seems to play a role in bone homeostasis (By
CC -!- similarity).
CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
CC -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF331844; AAK16158.1; -
CC EMBL; AF326736; AAK13451.1; -
CC EMBL; AF326739; AAK13454.1; -
CC Genew; HGNC:13771; SOST.
CC MIM; 605740; -
CC MIM; 269500; -
CC InterPro; IPR000359; Cys_knot.
CC SMART; SM00041; CT; 1.
CC PROSITE; PS01185; CTCK_1; FALSE_NEG.
CC PROSITE; PS01225; CTCK_2; FALSE_NEG.
CC Signal; Glycoprotein.
CC CHAIN 1 23 POTENTIAL.
CC DOMAIN 24 213 SCLEROSTIN.
CC FT CARBOHYD 82 172 CTCK.
CC FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SEQUENCE 213 AA; 23908 MW; 6DA7B5EDE674728A CRC64;

Query Match 17.3%; Score 32; DB 1; Length 213;
Best Local Similarity 100.0%; Pred. No. 1.6e-27; Indels 0; Gaps 0;
Matches 32; Conservative 0; Mismatches 0;

Qy 65 TDGCRSAKPVTGLVCSGCGCPARLLPNAIGR 96
Db 90 TDGCRSAKPVTGLVCSGCGCPARLLPNAIGR 121

RESULT 5

```

```

SOST_HUMAN STANDARD; PRT; 213 AA.
AC Q9BG78;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sclerostin precursor.
GN SOST.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=11181578;
RA Balemans W., Ebeling M., Patel N., van Hul E., Olson P., Droszegi M.,
RA Iacarca C., Wuyts W., van den Ende J., Willems P., Paes-Alves A.F.,
RA Hill S., Bueno M., Ramos F.J., Tacconi P., Dikkers F.G., Stratakis C.,
RA Lindpaintner K., Vickery B., Foerzler D., Van Hul W.;
RT "Increased bone density in sclerosteosis is due to the deficiency of a
RT novel secreted protein (SOST).";
RL Hum. Mol. Genet. 10:537-543(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX PubMed=11179006;
RA Brunkow M.E., Gardner J.C., Van Ness J., Paepser B.W., Kovacevich B.R.,
RA Prohl S., Skonier J.E., Zhao L., Sabo P.J., Fu Y.H., Alisch R.S.,
RA Gillett L., Colbert T., Tacconi P., Galas D., Hamersma H.,
RA Beighton P., Mulligan J.T.;
RA "Bone dysplasia sclerosteosis results from loss of the SOST gene
RT product, a novel cysteine knot-containing protein.";
RL Am. J. Hum. Genet. 68:577-589(2001).
CC -!- FUNCTION: Seems to play a role in bone homeostasis.
CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
CC -!- TISSUE SPECIFICITY: Widely expressed at low levels with highest
CC levels in bone, cartilage, kidney, liver, bone marrow and primary
CC osteoblasts differentiated for 21 days.
CC -!- DISEASE: Defects in SOST are the cause of sclerosteosis, a
CC progressive sclerosing bone dysplasia.
CC -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF331844; AAK16158.1; -
CC EMBL; AF326736; AAK13451.1; -
CC EMBL; AF326739; AAK13454.1; -
CC Genew; HGNC:13771; SOST.
CC MIM; 605740; -
CC MIM; 269500; -
CC InterPro; IPR000359; Cys_knot.
CC SMART; SM00041; CT; 1.
CC PROSITE; PS01185; CTCK_1; FALSE_NEG.
CC PROSITE; PS01225; CTCK_2; FALSE_NEG.
CC Signal; Glycoprotein.
CC CHAIN 1 23 POTENTIAL.
CC DOMAIN 24 213 SCLEROSTIN.
CC FT CARBOHYD 82 172 CTCK.
CC FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SEQUENCE 213 AA; 24030 MW; 30BBD55CE73D5BB2 CRC64;

Query Match 17.3%; Score 32; DB 1; Length 213;
Best Local Similarity 100.0%; Pred. No. 1.6e-27; Indels 0; Gaps 0;
Matches 32; Conservative 0; Mismatches 0;

Qy 65 TDGCRSAKPVTGLVCSGCGCPARLLPNAIGR 96
Db 90 TDGCRSAKPVTGLVCSGCGCPARLLPNAIGR 121

```


use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; AF326741; AAK13455.1; -
EMBL; AK017295; BAB30678.1; -
EMBL; AF326737; AAK13452.1; -
MGD; MGI:1921749; Soet.
DR InterPro; IPR000359; Cys_knot.
DR PROSITE; PS01185; CTCK_1; FALSE_NEG.
DR PROSITE; PS01225; CTCK_2; FALSE_NEG.
KW Signal; Glycoprotein.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 211 SCLEROSTIN.
FT DOMAIN 80 170 CTCK.
FT CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 173 173 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 72 72 G -> D (IN REF. 2).
SQ SEQUENCE 211 AA; 23443 MW; AEB094E358E34961 CRC64;

Query Match 67.6%; Score 125; DB 1; Length 211;
Best Local Similarity 100.0%; Pred. No. 6e-130;
Matches 125; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 50 VSEYSCRELHYTRLTDPGCRSAKPVTGLVCSGCGPARLLPNAIGRVKWRPNPGDFRC 109
DB 73 VSEYSCRELHYTRLTDPGCRSAKPVTGLVCSGCGPARLLPNAIGRVKWRPNPGDFRC 132
QY 110 IPDYYAQRVQLLCPGGAAPRSRKRLVASCCKKRLTRFHNOSELKDFGPETARPOKGRK 169
DB 133 IPDYYAQRVQLLCPGGAAPRSRKRLVASCCKKRLTRFHNOSELKDFGPETARPOKGRK 192
QY 170 PRGGA 174
DB 193 PRGGA 197

RESULT 2
SOST RAT
ID -SOST RAT STANDARD; PRT; 213 AA.
AC Q99F67;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sclerostin precursor.
GN SOST.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=21090529; PubMed=11179006;
RA Brunkow M.E., Gardner J.C., Van Ness J., Paepers B.W., Kovacevich B.R.,
RA Proll S., Skonier J.E., Zhao L., Sabo P.J., Fu Y.H., Alisch R.S.,
RA Gillett L., Colbert T., Tacconi P., Galas D., Hamersma H.,
RA Beighton P., Mulligan J.T.;
RT "Bone dysplasia sclerosteosis results from loss of the SOST gene product, a novel cysteine knot-containing protein.";
RL Am. J. Hum. Genet. 68:577-589(2001).
CC -!- FUNCTION: Seems to play a role in bone homeostasis (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
CC -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

or send an email to license@isb-sib.ch).

EMBL; AF326741; AAK13456.1; -
InterPro; IPR000359; Cys_knot.
DR PROSITE; PS01185; CTCK_1; FALSE_NEG.
DR PROSITE; PS01225; CTCK_2; FALSE_NEG.
KW Signal; Glycoprotein.
FT SIGNAL 1 28 POTENTIAL.
FT CHAIN 29 213 SCLEROSTIN.
FT DOMAIN 82 172 CTCK.
FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 213 AA; 23974 MW; 6C56C878BCBCD684B CRC64;

Query Match 58.4%; Score 108; DB 1; Length 213;
Best Local Similarity 100.0%; Pred. No. 3.2e-111;
Matches 108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 65 TDGFCRSAPVTGLVCSGCGPARLLPNAIGRVKWRPNPGDFRCIPDRYRAQRVQLLCP 124
DB 90 TDGFCRSAPVTGLVCSGCGPARLLPNAIGRVKWRPNPGDFRCIPDRYRAQRVQLLCP 149
QY 125 GGAAPRSRKRLVASCCKKRLTRFHNOSELKDFGPETARPOKGRKRP 172
DB 150 GGAAPRSRKRLVASCCKKRLTRFHNOSELKDFGPETARPOKGRKRP 197

RESULT 3
SOST BOVIN
ID -SOST BOVIN STANDARD; PRT; 176 AA.
AC Q9BG79;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sclerostin (Fragment).
GN SOST.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21090529; PubMed=11179006;
RA Brunkow M.E., Gardner J.C., Van Ness J., Paepers B.W., Kovacevich B.R.,
RA Proll S., Skonier J.E., Zhao L., Sabo P.J., Fu Y.H., Alisch R.S.,
RA Gillett L., Colbert T., Tacconi P., Galas D., Hamersma H.,
RA Beighton P., Mulligan J.T.;
RT "Bone dysplasia sclerosteosis results from loss of the SOST gene product, a novel cysteine knot-containing protein.";
RL Am. J. Hum. Genet. 68:577-589(2001).
CC -!- FUNCTION: Seems to play a role in bone homeostasis (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
CC -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; AF326738; AAK13453.1; -
InterPro; IPR000359; Cys_knot.
DR SMART; SM00041; CT; 1
DR PROSITE; PS01185; CTCK_1; FALSE_NEG.
DR PROSITE; PS01225; CTCK_2; FALSE_NEG.
KW Glycoprotein.
FT SIGNAL 1 1
FT DOMAIN 51 141 CTCK.
FT CARBOHYD 22 22 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:18:57 ; Search time 7.89333 Seconds
(without alignments)
972.101 Million cell updates/sec

Title: US-09-867-274-4
Perfect score: 185
Sequence: 1 QGQAFNDATVIFGLGEY.....KGRKPRGAKANQALENAY 185

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size : 0

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	DB ID	Description
1	125	67.6	211	1 SOST MOUSE	Q99p68 mus musculus
2	108	58.4	213	1 SOST RAT	Q99p67 rattus norv
3	32	17.3	176	1 SOST BOVIN	Q9bg79 bos taurus
4	32	17.3	213	1 SOST CERAE	Q9bg78 cercopithec
5	32	17.3	213	1 SOST HUMAN	Q9bg74 homo sapien
6	8	4.3	428	1 ELK1_HUMAN	P19419 mus musculu
7	8	4.3	429	1 ELK1_MOUSE	P19469 mus musculu
8	7	3.8	97	1 VE7_HPV44	Q80914 human papil
9	7	3.8	110	1 RL22_LEPIN	Q9xd31 leptospira
10	7	3.8	535	1 DPA_ECOLI	P23847 escherichia
11	7	3.8	632	1 VGLG_SVNV	P27277 sonchus yel
12	7	3.8	852	1 CSG_HAHLA	P08198 halobacteri
13	7	3.8	1168	1 MYSC_ACACA	P10569 acanthamoeb
14	7	3.8	1325	1 BCC3_ACEXY	Q9wx63 acetobactria
15	6	3.2	49	1 TX25_PHONI	P29424 phonetria
16	6	3.2	107	1 TH12_CORNE	P07887 corynebacte
17	6	3.2	110	1 RL22_ECOLI	P02423 escherichia
18	6	3.2	111	1 RL22_ACHLA	P29222 acholeplasma
19	6	3.2	112	1 RL22_SPICI	O31160 spiroplasma
20	6	3.2	113	1 RL22_TRETH	P48286 thermus the
21	6	3.2	115	1 TIAF_HUMAN	O95411 homo sapien
22	6	3.2	115	1 TIAF_MOUSE	Q92184 mus musculu
23	6	3.2	118	1 REV_HVILW	Q96224 human immun
24	6	3.2	121	1 RL7_LACLA	Q9c942 lactococcus
25	6	3.2	131	1 RL22_PHYS1	O66094 phytoplasma
26	6	3.2	166	1 RS10_ICTPU	Q90vr4 ictalurus p
27	6	3.2	167	1 B3AR_MERUN	O70432 meriones un
28	6	3.2	173	1 FRIS_LYNST	P42577 lymnaea eta
29	6	3.2	182	1 YCYO_YEAST	P25654 saccharomyc
30	6	3.2	193	1 HS72_CANAL	P46587 candida alb
31	6	3.2	195	1 YL47_ARCFU	O28135 archaeoglob
32	6	3.2	196	1 RETB_CHICK	P41263 gallus gall
33	6	3.2	198	1 RECR_STRPY	Q99z33 streptococc

P13997 kluyveromyc
P43855 haemophilus
Q9hx21 pseudomonas
O60258 homo sapien
O70627 mus musculu
Q8xxf9 ralstonia s
P01236 homo sapien
P55151 macaca mula
O47428 brachiosteo
O79417 branchiosteo
P48534 pismus sativ
Q9zeal rickettsia

ALIGNMENTS

RESULT 1
SOST_MOUSE STANDARD; PRT; 211 AA.

AC Q99p68; Q9D3L7;
15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sclerostin precursor.

GN SOST.

OS Mus musculus (Mouse)..
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

RN [1]
SEQUENCE FROM N.A.

RC STRAIN=129/SVJ;
MEDLINE=21090529; PubMed=11179006;

RA Brunkow M.E., Gardner J.C., Van Ness J., Paepers B.W., Kovacevich B.R.,
RA Proll S., Skonier J.E., Zhao L., Sabo P.J., Fu Y.H., Alisch R.S.,
RA Gillett L., Colbert T., Tacconi P., Galas D., Hamersma H.,
RA Beighton P., Mulligan J.T.,
RA "bone dysplasia sclerosteosis results from loss of the SOST gene
RT product, a novel cysteine knot-containing protein.";

RL Am. J. Hum. Genet. 68:577-589 (2001).

RN [2]
SEQUENCE FROM N.A.

RP STRAIN=C57BL/6J; TISSUE=Head;

RC STRAIN=21085660; PubMed=11217851;

RK Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stauber P., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barah G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hall D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Sakamoto N.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schenbach C., Seva T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayaishizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690 (2001).

CC -1- FUNCTION: Seems to play a role in bone homeostasis (By
similarity).

CC -1- SUBCELLULAR LOCATION: Secreted (potential).

CC -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTEINE KNOT-LIKE (CTCK) DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its

A:Gene: UU417

A:Genetic code: SGC3

Query Match

Best Local Similarity 3.8%; Score 7; DB 2; Length 321;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 KVLVLAS 139

|||||

DB 276 KVLVLAS 282

RESULT 12

S72647

Hypothetical protein lb - anthracnose fungus (Colletotrichum gloeosporioides) retrotrans

C:Species: Glomerella cingulata, Colletotrichum gloeosporioides

C:Date: 24-Oct-1998 #sequence_revision 24-Oct-1998 #text_change 17-Mar-1999

C:Accession: S72647

R:He, C.; Nourse, J.P.; Kelemu, S.; Irwin, J.A.G.; Manners, J.M.

Mol. Gen. Genet. 252, 320-331, 1996

A:Title: Cgtr1: a non-LTR retrotransposon with restricted distribution in the fungal phy

A:Reference number: 372619; MUID:96439839; PMID:8842152

A:Accession: S72647

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-328 <HEH>

A:Experimental source: biotype B, isolate UQ62

C:Genetics:

A:Mobile element: retrotransposon Cgtr1

Query Match

Best Local Similarity 3.8%; Score 7; DB 2; Length 328;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 85 GPARLLP 91

|||||

DB 123 GPARLLP 129

RESULT 13

C83577

Hypothetical protein PA0549 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa

C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C:Accession: C83577

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,

; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A:Reference number: A82950; MUID:20437337; PMID:10984043

A:Accession: C83577

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-354 <STO>

A:Cross-references: GB:AE004491; GB:AE004091; MUID:99946412; PIDN:AA03938.1; GSPDB:GN001

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA0549

Query Match

Best Local Similarity 3.8%; Score 7; DB 2; Length 354;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 95 GRVKWR 101

|||||

DB 206 GRVKWR 212

RESULT 14

S24802

polyferredoxin 6x2[4Fe-4S] vhuB [similarity] - Methanococcus voltae

C:Species: Methanococcus voltae

C:Date: 10-Sep-1999 #sequence_revision 18-Aug-2000 #text_change 19-Jan-2001

C:Accession: S24802

R:Halboth, S.; Klein, A.

Submitted to the EMBL Data Library, August 1991

A:Description: Methanococcus voltae harbors two gene groups each of homologous (Nife) - ar

A:Reference number: S16721

A:Accession: S24802

A:Molecule type: DNA

A:Residues: 1-398 <HAL>

A:Cross-references: EMBL:X61204; MUID:91747406; PIDN:CAA43512.1; PID:G1747410

A:Experimental source: strain PS(DSM1537)

R:Halboth, S.; Klein, A.

Mol. Gen. Genet. 233, 217-224, 1992

A:Title: Methanococcus voltae harbors four gene clusters potentially encoding two [Nife]

A:Reference number: A59304; MUID:92233118; PMID:1603063

A:Contents: annotation

C:Genetics:

A:Gene: vhuB

C:Superfamily: polyferredoxin 6x2[4Fe-4S]; ferredoxin 2[4Fe-4S] homology

F:4-52/Domain: ferredoxin 2[4Fe-4S] homology <FER1>

F:56-109/Domain: ferredoxin 2[4Fe-4S] homology <FER2>

F:125-179/Domain: ferredoxin 2[4Fe-4S] homology <FER3>

F:192-247/Domain: ferredoxin 2[4Fe-4S] homology <FER4>

F:261-329/Domain: ferredoxin 2[4Fe-4S] homology <FER5>

F:341-395/Domain: ferredoxin 2[4Fe-4S] homology <FER6>

Query Match

Best Local Similarity 3.8%; Score 7; DB 1; Length 398;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 SAKPVTE 77

|||||

DB 252 SAKPVTE 258

RESULT 15

C89753

protein FlilC7.3 [imported] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Nov-2001

C:Accession: C89753

R:Anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology

A:Reference number: A75000; MUID:99069613; PMID:9851916

A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.eleg

A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A:Accession: C89753

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-425 <STO>

A:Cross-references: GB:chr_X; PIDN:AAC69015.1; PID:g2914120; GSPDB:GN00028; CESP:FlilC7.3

C:Genetics:

A:Gene: FlilC7.3

A:Map position: X

Query Match

Best Local Similarity 3.8%; Score 7; DB 2; Length 425;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 152 SELKDFG 158

|||||

DB 332 SELKDFG 338

Search completed: March 28, 2003, 14:23:59

Job time : 16.3067 secs

QY 172 PGAKANQ 178

Db 6 PGAKANQ 12

RESULT 7

G81057 hypothetical protein NMB1656 [imported] - Neisseria meningitidis (strain MC58 serogroup

C;Species: Neisseria meningitidis

C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001

C;Accession: G81057

R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.

Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;

Xi, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Massignani, V.; Pizzi, M.

Science 287, 1809-1815, 2000

A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; V

A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A;Reference number: A81000; MUID:20175755; PMID:10710307

A;Accession: G81057

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-197 <TET>

A;Cross-references: GB:AE002516; GB:AE002098; NID:g7226905; PIDN:AAF42005.1; PID:g722690

A;Experimental source: serogroup B, strain MC58

C;Genetics:

A;Gene: NMB1656

QY 175 KANQAE 181

Db 122 KANQAE 128

RESULT 8

A81819 hypothetical protein NMA1913 [imported] - Neisseria meningitidis (strain Z2491 serogroup

C;Species: Neisseria meningitidis

C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001

C;Accession: A81819

R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel

; Holroyd, S.; Jagsels, K.; Leather, S.; Mungall, K.; Quail, M.A.; Rajandream,

Nature 404, 502-506, 2000

A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.

A;Reference number: A81775; MUID:20222556; PMID:10761919

A;Accession: A81819

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-197 <PAR>

A;Cross-references: GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CAB85134.1; PID:g738054

A;Experimental source: serogroup A, strain Z2491

C;Genetics:

A;Gene: NMA1913

QY 175 KANQAE 181

Db 122 KANQAE 128

RESULT 9

AI2930 hypothetical protein Atu3047 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)

C;Species: Agrobacterium tumefaciens

C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002

C;Accession: AI2930

R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell

; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, I

ster, E.W.

A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A;Reference number: AB2577; PMID:11743193

A;Accession: AI2930

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-296 <KUR>

A;Cross-references: GB:AE008689; PIDN:AA43863.1; PID:g17741408; GSPDB:GN00187

A;Experimental source: strain C58 (Dupont)

C;Genetics:

A;Gene: Atu3047

A;Map position: linear chromosome

C;Superfamily: oligopeptide permease protein oppB

Query Match 3.8%; Score 7; DB 2; Length 296;

Best Local Similarity 100.0%; Pred. No. 25;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 89 LLPNAIG 95

Db 206 LLPNAIG 212

RESULT 10

E98351 oligopeptide transport system permease protein appC AGR_L_3516 [imported] - Agrobacterium

C;Species: Agrobacterium tumefaciens

C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002

C;Accession: E98351

R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorillo, B.; Goldman,

A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;

Science 294, 2323-2328, 2001

A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A;Reference number: A97359; PMID:11743194

A;Accession: E98351

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-311 <KUR>

A;Cross-references: GB:AE007870; PIDN:AAK90335.1; PID:g15160372; GSPDB:GN00170

C;Genetics:

A;Gene: AGR_L_3516

A;Map position: linear chromosome

C;Superfamily: oligopeptide permease protein oppB

Query Match 3.8%; Score 7; DB 2; Length 311;

Best Local Similarity 100.0%; Pred. No. 27;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 89 LLPNAIG 95

Db 221 LLPNAIG 227

RESULT 11

B82892 conserved hypothetical UUA17 [imported] - Ureaplasma urealyticum

C;Species: Ureaplasma urealyticum

C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000

C;Accession: B82892

R;Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.

submitted to GenBank, February 2000

A;Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mini

A;Reference number: A82870

A;Accession: B82892

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-321 <GLA>

A;Cross-references: GB:AE002139; GB:AF222894; NID:56899405; PIDN:AAF30828.1; GSPDB:GN001

A;Experimental source: serovar 3; biovar 1

C;Genetics:

A;Cross-references: EMBL:X87257; NID:g836634; PIDN:CAA60715.1; PID:g836635
A;Experimental source: embryo
R;Giovane, A.; Pintzas, A.; Maira, S.M.; Sobiesczuk, P.; Wasyluk, B.
Genes Dev. 8, 1502-1513, 1994
A;Title: Net, a new ets transcription factor that is activated by Ras.
A;Reference number: A53837; MUID:95047310; PMID:7958835
A;Accession: I48339
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-429 <RES>
A;Cross-references: EMBL:X87257; NID:g836634; PIDN:CAA60715.1; PID:g836635
A;Accession: I48340
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 5-132, "T", 134-224 <RE2>
A;Cross-references: EMBL:Z36939; NID:g535922; PIDN:CAA85391.1; PID:g535923
C;Comment: This protein belongs to the subfamily of ternary complex factor (TCF) which b
C;Genetics:
A;Gene: elk1
A;Introns: 70/3, 219/3, 363/3, 397/3
C;Superfamily: elk-1 transforming protein; ets DNA-binding domain homology
F;7-86/Domain: ets DNA-binding domain homology <ETS>

Query Match 4.3%; Score 8; DB 2; Length 429;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 PQGRKPR 171
|||||
DB 311 PQGRKPR 318

RESULT 3
AG0002
conserved hypothetical protein YPO0013a [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C;Accession: AG0002
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
Li, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AG0002
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-89 <KUR>
A;Cross-references: GB:AL590842; PIDN:CAC88881.1; PID:g15978129; GSPDB:GN00175
C;Genetics:
A;Gene: YPO0013a

Query Match 3.8%; Score 7; DB 2; Length 89;
Best Local Similarity 100.0%; Pred. No. 9.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 TEVIPGL 17
|||||
DB 62 TEVIPGL 68

RESULT 4
D83305
hypothetical protein PA2722 [imported] - Pseudomonas aeruginosa (strain PA01)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: D83305
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: AB2950; MUID:20437337; PMID:10984043

A;Accession: D83305
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-130 <STO>
A;Cross-references: GB:AE004700; GB:AE004091; NID:g9948792; PIDN:AAG06110.1; GSPDB:GN001;
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA2722

Query Match 3.8%; Score 7; DB 2; Length 130;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 KVLVAS 139
|||||
DB 12 KVLVAS 18

RESULT 5
CB3091
hypothetical protein PA4441 [imported] - Pseudomonas aeruginosa (strain PA01)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: CB3091
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: AB2950; MUID:20437337; PMID:10984043
A;Accession: CB3091
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-148 <STO>
A;Cross-references: GB:AE004858; GB:AE004091; NID:g9950668; PIDN:AAG07829.1; GSPDB:GN001;
A;Experimental source: strain PA01
C;Genetics:
A;Gene: PA4441

Query Match 3.8%; Score 7; DB 2; Length 148;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 87 ARLLPNA 93
|||||
DB 26 ARLLPNA 32

RESULT 6
AC1832
hypothetical protein alr0203 [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp.
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C;Accession: AC1832
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yanada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anat
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AC1832
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-162 <KUR>
A;Cross-references: GB:BA000019; PIDN:BA07727.1; PID:g17135181; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: alr0203

Query Match 3.8%; Score 7; DB 2; Length 162;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:20:22 ; Search time 14.3067 Seconds
(without alignments)
1243.118 Million cell updates/sec

Title: US-09-867-274-4
Perfect score: 185
Sequence: 1 QGQAFRNDATVPIGLGEY.....KGRKPRPCAKAQAELNAY 195

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR 73.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	4.3	428	1 TVHUEK	transforming prote
2	8	4.3	429	2 JC4965	elk1 protein - mou
3	7	3.8	89	2 AG0002	conserved hypothet
4	7	3.8	130	2 D83305	hypothetical prote
5	7	3.8	148	2 C83091	hypothetical prote
6	7	3.8	162	2 AC1832	hypothetical prote
7	7	3.8	197	2 G81057	hypothetical prote
8	7	3.8	197	2 A81819	hypothetical prote
9	7	3.8	296	2 AI2930	hypothetical prote
10	7	3.8	311	2 E98351	oligopeptide trans
11	7	3.8	321	2 B82892	conserved hypothet
12	7	3.8	328	2 S72647	hypothetical prote
13	7	3.8	354	2 C93577	hypothetical prote
14	7	3.8	398	1 S24802	polyferredoxin 6x2
15	7	3.8	425	2 C89753	protein Fl1C7.3 [i
16	7	3.8	531	2 B83082	probable binding p
17	7	3.8	535	1 A39194	periplasmic dipept
18	7	3.8	535	2 H91181	dipeptide transpor
19	7	3.8	535	2 D86028	dipeptide dipept
20	7	3.8	535	2 A80983	periplasmic dipept
21	7	3.8	542	2 T23869	hypothetical prote
22	7	3.8	562	2 T05758	hypothetical prote
23	7	3.8	565	2 B70652	probable cyto prot
24	7	3.8	632	1 VGVNSY	surface glycoprote
25	7	3.8	632	2 AG1547	transcription anti
26	7	3.8	681	2 E82812	outer membrane hem
27	7	3.8	730	2 B87266	TonB-dependent rec
28	7	3.8	836	2 B84417	cell surface glyco
29	7	3.8	852	2 A28459	cell surface glyco

probable enzyme wi
myosin heavy chain
hypothetical prote
polyketide synthas
myosin heavy chain
ribosomal protein
neurotoxin Tx2 - g
hypothetical prote
unknown protein en
hypothetical prote
hypothetical prote
thioredoxin C-2 -
hypothetical prote
hypothetical prote
ribosomal protein
50S ribosomal subu

30 7 3.8 1008 2 B95979
31 7 3.8 1168 1 MWAXIC
32 7 3.8 1308 2 T15280
33 7 3.8 1488 2 AG2136
34 7 3.8 1940 2 A59287
35 6 3.2 22 2 JF0066
36 6 3.2 49 2 S29215
37 6 3.2 89 2 D75271
38 6 3.2 102 2 G85584
39 6 3.2 102 2 E90734
40 6 3.2 103 2 G84293
41 6 3.2 108 1 S02802
42 6 3.2 108 2 F82480
43 6 3.2 108 2 D90051
44 6 3.2 110 1 R5EC22
45 6 3.2 110 2 D91151

ALIGNMENTS

transforming protein elk-1 - human
C:Species: Homo sapiens (man)
C:Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Jun-1999
C:Accession: A41354; S54721
R:Rao, V.N.; Huebner, K.; Isobe, M.; ar-Rushdi, A.; Croce, C.M.; Reddy, E.S.P.
Science 244, 66-70, 1989
A>Title: elk, tissue-specific ets-related genes on chromosomes X and 14 near translocation
A:Reference number: A41354; MUID:89203250; PMID:2539641
A:Accession: A41354
A:Molecule type: mRNA
A:Residues: 1-428 <RAO>
A:Cross-references: GB:M25269; NID:9538208; PIDN:AAA52384.1; PID:9538209
R:Gille, H.; Kortenjann, M.; Thomae, O.; Moomaw, C.; Slaughter, C.; Cobb, M.H.; Shaw, P.I
EMBO J. 14, 951-962, 1995
A>Title: ERK phosphorylation potentiates Elk-1-mediated ternary complex formation and tra
A:Reference number: S54721; MUID:95196758; PMID:7889942
A:Accession: S54721
A>Status: preliminary
A:Molecule type: protein
A:Residues: 318-328, 'XX', 331, 336-364; 380-388, 'X', 390-392, 'X', 394-400, 'XX', 403-405, 'X', 40:
C:Genetics:
A:Gene: GDB:ELK1
A:Cross-references: GDB:119867; OMIM:311040
A:Map position: Xp11.2-Xp11.2
C:Superfamily: elk-1 transforming protein; ets DNA-binding domain homology
C:Keywords: DNA binding; oncogene; transforming protein
F:7-86/Domain: ets DNA-binding domain homology <ETS>

Query Match 4.3%; Score 8; DB 1; Length 428;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 PQGRKPR 171
DB 310 PQGRKPR 317

RESULT 2
JC4965
C:Species: Mus musculus (house mouse)
C:Date: 31-Dec-1996 #sequence_revision 31-Dec-1996 #text_change 16-Jul-1999
C:Accession: JC4965; I48339; I48340; S54908
R:Grevin, D.; Ung, S.; Denhez, F.; Dehem, M.; Quatmanns, B.; Begue, A.; Stehelin, D.; Me
Gene 174, 185-188, 1996
A>Title: Structure and organization of the mouse elk1 gene.
A:Reference number: JC4965; MUID:97017146; PMID:8863747
A:Accession: JC4965
A:Molecule type: mRNA
A:Residues: 1-429 <GRE>

Mon Mar 31 09:44:32 2003

```

PR 28-FEB-2001; 2001WO-US06520.
PR 09-MAR-2001; 2001US-274399P.
PR 03-APR-2001; 2001US-280982P.
PR 04-APR-2001; 2001US-282129P.
PR 04-APR-2001; 2001US-282199P.
PR 09-MAY-2001; 2001US-290589P.
PR 25-MAY-2001; 2001WO-US17092.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
XX
XX (GETH ) GENENTECH INC.
XX
XX Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC;
PI Gurney AL, Smith V, Stephan J, Watanabe CK, Wood WI, Zhang Z;
PI Fong S;
XX
XX WPI; 2002-362426/39.
DR N-PSDB; ABK69992.
XX
XX New PRO polypeptides and polynucleotides encoding the polypeptides,
PT useful in gene therapy, chromosome identification, tissue typing, or
PT for genetic analysis of individuals with genetic disorders
XX
XX Claim 11; Figure 64; 218pp; English.
XX
XX This invention relates to the cDNA and protein sequences of novel
CC secreted and transmembrane polypeptides PRO polypeptides. The
CC invention also comprises a method for producing the proteins of the
CC invention by recombinant means and antibodies specific for the protein
CC of the invention. The antibody may be used for detecting the PRO
CC proteins of the invention and may be used to modify their activity.
CC polynucleotides may be used as hybridisation probes for a cDNA library
CC to isolate the full-length PRO cDNA or to isolate other cDNAs, to
CC construct hybridisation probes for mapping the gene which encodes that
CC PRO and for genetic analysis of individuals with genetic disorders, in
CC assays to identify other proteins or molecules involved in binding
CC reaction, to generate transgenic animals or knock-out animals which in
CC turn are useful in the development and screening of therapeutically
CC useful reagents, for chromosome identification, and tissue typing. The
CC PRO polypeptides are useful in gene therapy, and as molecular weight
CC markers for protein electrophoresis purposes. The sequences may
CC also be used to detect overexpression on PRO polypeptides in cancerous
CC tumours and for screening for differentially expressed genes using
CC microarray technology. The present sequence represents a human PRO
CC protein of the invention.
XX
XX Sequence 213 AA;
SQ
Query Match 17.3%; Score 32; DB 23; Length 213;
Best Local Similarity 100.0%; Pred. No. 7.5e-25;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 65 TDGFCRSAPVTELVCSCGCGPARLLPNAIGR 96
DB 90 TDGFCRSAPVTELVCSCGCGPARLLPNAIGR 121

```

Search completed: March 28, 2003, 14:21:58
 Job time : 33.56 secs

DR WPI; 2000-412321/35.
 XX N-PSDB; AAA29062.
 PT Nucleic acids (I) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures
 XX
 PS Disclosure; Page 121; 162pp; English.
 XX
 CC This shows a variant human transforming growth factor-beta
 CC (TGF-beta) binding protein designated BEER P38R. The encoded protein
 CC comprises a substitution of arginine for the wild-type proline at
 CC residue 38. The cDNA and protein may be used for prevention, treatment
 CC and diagnosis of diseases associated with inappropriate BEER expression.
 CC For example, they may be used to treat disorders associated with
 CC decreased TGF-beta BP expression. The cDNA or vectors may be administered
 CC to treat diseases by rectifying mutations or deletions in a patient's
 CC genome that affect the activity of BEER by expressing inactive proteins
 CC or to supplement the patient's own production of BEER polypeptides. The
 CC nucleic acids may be used for recombinant production of BEER. Gene
 CC therapy, antisense therapy, as probes for diagnostic assays and for
 CC functional studies. BEER may be used to raise antibodies and for
 CC identification of BEER modulators. BEER antagonists may be used to
 CC increase bone mineral content for the treatment of disorders such as
 CC osteopenia, osteoporosis, fractures and other disorders associated with
 CC low mineral content.
 XX
 SQ Sequence 213 AA;
 Query Match 17.3%; Score 32; DB 21; Length 213;
 Best Local Similarity 100.0%; Pred. No. 7.5e-25;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 65 TDGFCRSAPVTELVCSGCGPARLLPNAIGR 96
 DB 90 TDGFCRSAPVTELVCSGCGPARLLPNAIGR 121
 RESULT 14
 AAAY97589
 ID RAY97589 standard; Protein; 213 AA.
 AC RAY97589;
 XX
 DT 05-APR-2001 (first entry)
 DE Human secreted protein PRO7476.
 XX
 KW Secreted protein; human; PRO protein; neoplastic cell growth; tumour;
 KW proliferation; leukaemia; lymphoid malignancy; inflammatory disorder;
 KW angiogenic disorder; immunologic disorder; PRO7476.
 XX
 OS Homo sapiens.
 XX
 PN WO200075317-A2.
 PD 14-DEC-2000.
 XX
 PF 15-MAY-2000; 2000WO-US13358.
 XX
 PR 09-JUN-1999; 99US-0138385.
 PR 20-JUL-1999; 99US-0144790.
 PR 03-AUG-1999; 99US-0146843.
 PR 10-AUG-1999; 99US-0148188.
 PR 17-AUG-1999; 99US-0149320.
 PR 17-AUG-1999; 99US-0149327.
 PR 17-AUG-1999; 99US-0149396.
 PR 20-AUG-1999; 99US-0150114.
 PR 31-AUG-1999; 99US-0151700.
 PR 31-AUG-1999; 99US-0151734.
 XX
 PA (GETH) GENENTECH INC.
 XX

PI Botstein DA, Goddard A, Gurney AL, Smith V, Watanabe CK, Wood WI;
 XX WPI; 2001-071075/08.
 DR N-PSDB; AAA91023.
 XX
 PT Antibodies against PRO polypeptides, useful for diagnosing and treating
 PT tumours are associated with gene amplification, neoplastic cell growth
 PT and proliferation in mammals -
 XX
 PS Claim 61; Fig 20; 143pp; English.
 XX
 CC This sequence is a human PRO protein of the invention. The PRO
 CC proteins are secreted proteins. Antagonists or antibodies of PRO
 CC polypeptides are useful for diagnosing and treating tumours are
 CC associated with gene amplification, neoplastic cell growth and
 CC proliferation in mammals, and those conditions characterised by
 CC overexpression and/or activation of the amplified genes. Such conditions
 CC include benign or malignant tumours (e.g. renal, liver, kidney, bladder,
 CC breast, gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,
 CC thyroid, hepatic carcinomas, sarcomas, glioblastomas and various head and
 CC neck tumours); leukemias and lymphoid malignancies; neuronal, glial
 CC astrocytic, hypothalamic, and other glandular, macrophageal, epithelial,
 CC stromal and blastoclastic disorders; and inflammatory, angiogenic and
 CC immunologic disorders. These may further be used to qualitatively or
 CC quantitatively detect the expression of proteins encoded by the
 CC amplified genes, and in tumour diagnostics or prognostics. The PRO
 CC polypeptide or its antagonist may be used for the preparation of a
 CC medicament in the treatment of a condition, which is responsive to the
 CC PRO polypeptide, its antagonist or anti-PRO antibody.
 XX
 SQ Sequence 213 AA;
 Query Match 17.3%; Score 32; DB 22; Length 213;
 Best Local Similarity 100.0%; Pred. No. 7.5e-25;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 65 TDGFCRSAPVTELVCSGCGPARLLPNAIGR 96
 DB 90 TDGFCRSAPVTELVCSGCGPARLLPNAIGR 121
 RESULT 15
 ABG34061
 ID ABG34061 standard; Protein; 213 AA.
 AC ABG34061;
 XX
 DT 15-JUL-2002 (first entry)
 DE Human Pro peptide #32.
 XX
 KW Human; PRO; secreted protein; transmembrane protein;
 KW genetic disorder; tumour; cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO200224888-A2.
 PD 28-MAR-2002.
 XX
 PF 29-AUG-2001; 2001WO-US27099.
 XX
 PR 01-SEP-2000; 2000US-239896P.
 PR 05-SEP-2000; 2000US-230621P.
 PR 22-SEP-2000; 2000US-235147P.
 PR 10-NOV-2000; 2000WO-US30873.
 PR 12-JAN-2001; 2001US-261878P.
 PR 16-JAN-2001; 2001US-261910P.
 PR 16-JAN-2001; 2001US-261939P.
 PR 25-JAN-2001; 2001US-262150P.
 PR 02-FEB-2001; 2001US-264395P.
 PR 09-FEB-2001; 2001US-266421P.
 PR 09-FEB-2001; 2001US-267623P.

Mon Mar 31 09:44:32 2003

```

XX PF 24-NOV-1999; 99WO-US27990.
XX PR 27-NOV-1998; 98US-0110283.
XX PA (DARW-) DARWIN DISCOVERY LTD.
XX PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepers BW;
XX PI Van Ness J, Winkler DG;
XX DR WPI; 2000-412321/35.
XX DR N-PSDB; AAA29056.
XX PT Nucleic acids (I) encoding a transforming growth factor beta binding
XX PT protein, useful for identifying agents for treating osteopenia,
XX PT osteoporosis and fractures
XX PS Claim 3; Page 119-120; 162pp; English.
XX CC This shows a variant human transforming growth factor-beta (TGF-beta)
XX CC binding protein designated BEER V101, which comprises a substitution of
XX CC isoleucine for the wild-type valine at residue 10. The cDNA and protein
XX CC may be used for prevention, treatment and diagnosis of diseases
XX CC associated with inappropriate BEER expression. For example, they may be
XX CC used to treat disorders associated with decreased TGF-beta BP expression.
XX CC The cDNA or vectors may be administered to treat diseases by rectifying
XX CC mutations or deletions in a patient's genome that affect the activity of
XX CC BEER by expressing inactive proteins or to supplement the patients own
XX CC production of BEER polypeptides. The nucleic acids may be used for
XX CC recombinant production of BEER, gene therapy, antisense therapy, as
XX CC probes for diagnostic assays and for functional studies. BEER may be used
XX CC to raise antibodies and for identification of BEER modulators. BEER
XX CC antagonists may be used to increase bone mineral content for the
XX CC treatment of disorders such as osteopenia, osteoporosis, fractures and
XX CC other disorders associated with low mineral content.
XX SQ Sequence 213 AA;
Query Match 17.3%; Score 32; DB 21; Length 213;
Best Local Similarity 100.0%; Pred. No. 7.5e-25;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 65 TDGPCRSAPVTTELVCSCGCGPARLLPNAIGR 96
Db 90 TDGPCRSAPVTTELVCSCGCGPARLLPNAIGR 121
RESULT 12
AAAY96431
ID AAAY96431 standard; Protein; 213 AA.
XX AC AAAY96431;
XX DT 12-SEP-2000 (first entry)
XX DE Vervet TGF-beta binding protein (BEER).
XX KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
XX KW BEER; gene therapy; antisense therapy; fracture; bone mineralization.
XX OS Cercopithecus pygerythrus.
XX PN WO200032773-A1.
XX PD 08-JUN-2000.
XX PF 24-NOV-1999; 99WO-US27990.
XX PR 27-NOV-1998; 98US-0110283.
XX PA (DARW-) DARWIN DISCOVERY LTD.
XX PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepers BW;
XX PI Van Ness J, Winkler DG;

```

```

PI Van Ness J, Winkler DG;
XX WPI; 2000-412321/35.
XX DR N-PSDB; AAA29057.
XX PT Nucleic acids (I) encoding a transforming growth factor beta binding
XX PT protein, useful for identifying agents for treating osteopenia,
XX PT osteoporosis and fractures
XX PS Claim 4; Page 122-123; 162pp; English.
XX CC This shows a vervet transforming growth factor-beta (TGF-beta)
XX CC binding protein designated VBEER. The cDNA and protein may be used for
XX CC prevention, treatment and diagnosis of diseases associated with
XX CC inappropriate BEER expression. For example, they may be used to treat
XX CC disorders associated with decreased TGF-beta BP expression. The cDNA or
XX CC vectors may be administered to treat diseases by rectifying mutations or
XX CC deletions in a patient's genome that affect the activity of BEER by
XX CC expressing inactive proteins or to supplement the patients own production
XX CC of BEER polypeptides. The nucleic acids may be used for recombinant
XX CC production of BEER, gene therapy, antisense therapy, as probes for
XX CC diagnostic assays and for functional studies. BEER may be used to raise
XX CC antibodies and for identification of BEER modulators. BEER antagonists
XX CC may be used to increase bone mineral content for the treatment of
XX CC disorders such as osteopenia, osteoporosis, fractures and other disorders
XX CC associated with low mineral content.
XX SQ Sequence 213 AA;
Query Match 17.3%; Score 32; DB 21; Length 213;
Best Local Similarity 100.0%; Pred. No. 7.5e-25;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 65 TDGPCRSAPVTTELVCSCGCGPARLLPNAIGR 96
Db 90 TDGPCRSAPVTTELVCSCGCGPARLLPNAIGR 121
RESULT 13
AAAY96436
ID AAAY96436 standard; Protein; 213 AA.
XX AC AAAY96436;
XX DT 12-SEP-2000 (first entry)
XX DE Human TGF-beta binding protein (BEER) variant P38R.
XX KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
XX KW BEER; variant; P38R; gene therapy; antisense therapy; fracture;
XX KW chromosome 17q12-21; bone mineralization.
XX OS Homo sapiens.
XX PH Key Location/Qualifiers
XX FT Misc-difference 38 /label= P38R
XX FT /note= "wild type proline has been substituted with
XX FT arginine"
XX PN WO200032773-A1.
XX PD 08-JUN-2000.
XX PF 24-NOV-1999; 99WO-US27990.
XX PR 27-NOV-1998; 98US-0110283.
XX PA (DARW-) DARWIN DISCOVERY LTD.
XX PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepers BW;
XX PI Van Ness J, Winkler DG;
XX

```

AAB26106
 ID AAB26106 standard; Protein; 213 AA.
 AC AAB26106;
 XX
 DT 15-JAN-2001 (first entry)
 XX
 DE Human DAN/Cerberus-related protein 6 (hDCR6) (exons 1 and 4).
 XX
 KW Human; DNA/Cerberus-related protein 6; hDCR6; morphogenic protein;
 KW antagonist; BMP; cell growth; cell differentiation; bone formation;
 KW gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO200055193-A2.
 XX
 PD 21-SEP-2000.
 XX
 PF 02-MAR-2000; 2000WO-US05537.
 XX
 PR 12-MAR-1999; 99US-0124118.
 XX
 PA (REGG-) REGENERON PHARM INC.
 XX
 PI Economides AN;
 XX
 DR WPI; 2000-638179/61.
 DR N-PSDB; AAA94051.
 XX
 FT Novel isolated, human DNA/Cerberus related protein 6 which include
 FT natural homologue, and polypeptides comprising DCR6 domain and nucleic
 FT acids encoding the proteins which are useful as probes and primers
 XX
 PS Claim 8; Fig 3; 40pp; English.
 XX
 CC The present sequence comprises the amino acid sequence encoded by exons 1
 CC and 4 of the human DAN/Cerberus-related protein 6 (hDCR6) coding
 CC sequence. The coding sequence was isolated from a human kidney cDNA
 CC library containing exons 1 and 4 of the sequence. hDCR6 is closely
 CC related to the DAN and DCR5 proteins, both of which act as antagonists of
 CC morphogenic proteins such as BMP. It is possible that the hDCR6 gene and
 CC protein can be used as immunogens, modulators of cell function, growth
 CC and differentiation, to reduce undesirable bone formation, to identify
 CC DCR6 binding agents, in diagnosis, and in gene therapy.
 XX
 SQ Sequence 213 AA;
 Query Match 17.3%; Score 32; DB 21; Length 213;
 Best Local Similarity 100.0%; Pred. No. 7.5e-25;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 65 TDGFCRSKPVTELVCSGCGPARLLPNAIGR 96
 Db |||||
 90 TDGFCRSKPVTELVCSGCGPARLLPNAIGR 121
 RESULT 10
 AAY96429
 ID AAY96429 standard; Protein; 213 AA.
 XX
 AC AAY96429;
 XX
 DT 12-SEP-2000 (first entry)
 XX
 DE Human TGF-beta binding protein (BEER).
 DE osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KW BEER; gene therapy; antisense therapy; fracture; bone mineralization.
 KW
 OS Homo sapiens.
 XX
 PN WO200032773-A1.
 XX
 PD

XX 08-JUN-2000.
 XX
 XX 24-NOV-1999; 99WO-US27990.
 XX
 XX 27-NOV-1998; 98US-0110283.
 XX
 XX (DARW-) DARWIN DISCOVERY LTD.
 XX
 PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepers BW;
 PI Van Ness J, Winkler DG;
 XX
 DR WPI; 2000-412321/35.
 DR N-PSDB; AAA29055.
 XX
 FT Nucleic acids (I) encoding a transforming growth factor beta binding
 FT protein, useful for identifying agents for treating osteopenia,
 FT osteoporosis and fractures
 XX
 PS Claim 2; Page 116; 152pp; English.
 XX
 CC This shows the human transforming growth factor-beta (TGF-beta)
 CC binding protein designated hBEER. The cDNA and protein may be used for
 CC prevention, treatment and diagnosis of diseases associated with
 CC inappropriate BEER expression. For example, they may be used to treat
 CC disorders associated with decreased TGF-beta BP expression. The cDNA or
 CC vectors may be administered to treat diseases by rectifying mutations or
 CC deletions in a patient's genome that affect the activity of BEER by
 CC expressing inactive proteins or to supplement the patient's own production
 CC of BEER polypeptides. The nucleic acids may be used for recombinant
 CC production of BEER, gene therapy, antisense therapy, as probes for
 CC diagnostic assays and for functional studies. BEER may be used to raise
 CC antibodies and for identification of BEER modulators. BEER antagonists
 CC may be used to increase bone mineral content for the treatment of
 CC disorders such as osteopenia, osteoporosis, fractures and other
 CC disorders associated with low mineral content.
 XX
 SQ Sequence 213 AA;
 Query Match 17.3%; Score 32; DB 21; Length 213;
 Best Local Similarity 100.0%; Pred. No. 7.5e-25;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 65 TDGFCRSKPVTELVCSGCGPARLLPNAIGR 96
 Db |||||
 90 TDGFCRSKPVTELVCSGCGPARLLPNAIGR 121
 RESULT 11
 AAY96430
 ID AAY96430 standard; Protein; 213 AA.
 XX
 AC AAY96430;
 XX
 DT 12-SEP-2000 (first entry)
 XX
 DE Human TGF-beta binding protein (BEER) variant V101.
 XX
 KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KW BEER; variant; V101; gene therapy; antisense therapy; fracture;
 KW chromosome 17q12-21; bone mineralization.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 10
 FT /label= V101
 FT /note= "wild type valine has been substituted with
 FT isoleucine"
 XX
 PN WO200032773-A1.
 XX
 PD 08-JUN-2000.

RESULT 7
 AAY96434
 ID AAY96434 standard; Protein; 176 AA.

XX AC AAY96434;
 XX DT 12-SEP-2000 (first entry)
 XX DE Bovine TGF-beta binding protein (BEER).
 XX DE osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KW BEER; gene therapy; antisense therapy; fracture; bone mineralization.
 XX OS Bos taurus.

XX PN WO200032773-A1.
 XX PD 08-JUN-2000.
 XX PF 24-NOV-1999; 99WO-US27990.
 XX PR 27-NOV-1998; 98US-0110283.
 XX PA (DARW-) DARWIN DISCOVERY LTD.

XX PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepker BW;
 PI Van Ness J, Winkler DG;
 DR WPI; 2000-412321/35.
 DR N-PSDB; AAA29060.

XX PT Nucleic acids (1) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures

XX PS Claim 7; Page 127; 162pp; English.

XX CC This shows a bovine transforming growth factor-beta (TGF-beta)
 CC binding protein designated bBEER. The cDNA and protein may be used for
 CC prevention, treatment and diagnosis of diseases associated with
 CC inappropriate BEER expression. For example, they may be used to treat
 CC disorders associated with decreased TGF-beta Bp expression. The cDNA or
 CC vectors may be administered to treat diseases by rectifying mutations or
 CC deletions in a patient's genome that affect the activity of BEER by
 CC expressing inactive proteins or to supplement the patients own production
 CC of BEER polypeptides. The nucleic acids may be used for recombinant
 CC production of BEER, gene therapy, antisense therapy, as probes for
 CC diagnostic assays and for functional studies. BEER may be used to raise
 CC antibodies and for identification of BEER modulators. BEER antagonists
 CC may be used to increase bone mineral content for the treatment of
 CC disorders such as osteopenia, osteoporosis, fractures and other disorders
 CC associated with low mineral content.

XX SQ Sequence 176 AA;
 Query Match 17.3%; Score 32; DB 21; Length 176;
 Best Local Similarity 100.0%; Pred. No. 6.4e-25;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 TDGFCRSAPVTELVCSGCGPARLLPNAIGR 96
 |||||
 DB 59 TDGFCRSAPVTELVCSGCGPARLLPNAIGR 90

RESULT 8
 ABB07207
 ID ABB07207 standard; Protein; 190 AA.
 XX AC ABB07207;
 XX DT 26-MAR-2002 (first entry)

XX DE Human cloaked-2 polypeptide mature protein sequence.

XX KW Cloaked-2; cysteine knot motif; nephrotropic; cardiant; immunomodulator;
 KW hepatotropic; antiinflammatory; antithyroid; cyrostatic; neuroprotective;
 KW antianemic; hypotensive; antiarrhythmic; antiarteriosclerotic; muscular;
 KW antidiabetic; anorectic; gene therapy; cell therapy; antisense therapy;
 human.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT Misc-difference 1..16
 FT /note= "this region has been repeated twice in the
 FT sequence provided in the sequence listing but has
 FT been indicated correctly in the sequence in the
 FT Figure"

XX PN WO200192308-A2.

XX PD 06-DEC-2001.

XX PF 29-MAY-2001; 2001WO-US17478.

XX PR 01-JUN-2000; 2000US-208550P.

XX PR 04-AUG-2000; 2000US-223542P.

XX PA (AMGE-) AMGEN INC.

XX PI Paszty CJ, Gao Y;

XX DR WPI; 2002-114325/15.

XX DR N-PSDB; ABA94293.

XX PT New human and mouse cystine-knot polypeptide designated as Cloaked-2,
 PT for treating or preventing kidney, heart (e.g. myocardial infarction)
 PT or liver (e.g. hepatitis) diseases -
 XX PS Claim 13; Fig 1; 170pp; English.

XX CC The invention relates to polypeptides comprising a cysteine knot motif
 CC and designated as Cloaked-2, derived from human and mouse. The cloaked-2
 CC polypeptides can be expressed by standard recombinant methodology. The
 CC cloaked-2 polynucleotides are useful in gene therapy and antisense
 CC therapy. The cloaked-2 polypeptides and polynucleotides are useful for
 CC treating, preventing, ameliorating or detecting diseases and disorders of
 CC the kidney (e.g. anemia, hypertension or low blood pressure), heart (e.g.
 CC cardiac hypertrophy, congestive heart failure, myocardial infarction,
 CC arrhythmias, atherosclerosis, hypertension or low blood pressure),
 CC skeletal muscle (e.g. muscular dystrophy or cachexia), placenta (e.g.
 CC congenital abnormalities or miscarriage), liver (e.g. hepatitis or
 CC cirrhosis), pancreas (e.g. diabetes or pancreatitis), thyroid (e.g.
 CC Grave's disease or myxedema) or adrenal cortex (e.g. Cushing's disease
 CC or Addison's disease), homeostasis or metabolic diseases (e.g. obesity,
 CC cancer or myopathies), infections, or autoimmune diseases. Selective
 CC binding agents may be used to modulate the biological activities of
 CC Cloaked-2 polypeptides or to detect Cloaked-2 polypeptide levels in a
 CC sample. Transgenic non-human animals are useful for drug candidate
 CC screening. The present sequence represents the human cloaked-2
 CC mature polypeptide sequence.

XX SQ Sequence 190 AA;

Query Match 17.3%; Score 32; DB 23; Length 190;
 Best Local Similarity 100.0%; Pred. No. 6.8e-25;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 TDGFCRSAPVTELVCSGCGPARLLPNAIGR 96

DB 67 TDGFCRSAPVTELVCSGCGPARLLPNAIGR 98

RESULT 9

XX Human; bone marrow expressed exon; gene expression analysis; probe;
 KW microarray; cancer; leukaemia; lymphoma; myeloma.
 XX Homo sapiens.
 XX WO200157276-A2.
 XX 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-US006668.
 XX 04-FEB-2000; 2000US-0180312.
 XX 26-MAY-2000; 2000US-0207456.
 XX 30-JUN-2000; 2000US-0608408.
 XX 03-AUG-2000; 2000US-0632366.
 XX 21-SEP-2000; 2000US-0234687.
 XX 27-SEP-2000; 2000US-0236359.
 XX 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-489900/53.
 XX Human genome-derived single exon nucleic acid probes useful for
 XX analyzing gene expression in human bone marrow -
 XX Example 4; SEQ ID NO: 33373; 689pp + Sequence Listing; English.
 XX The present invention provides a number of single exon nucleic acid
 XX probes which are derived from genomic sequences expressed in the human
 XX bone marrow. They can be used to measure gene expression in bone marrow
 XX samples, which may enable the improved diagnosis and treatment of cancers
 XX such as lymphoma, leukaemia and myeloma. The present sequence is a
 XX protein encoded by one of the probes of the invention.
 XX Sequence 139 AA;
 XX
 XX Query Match 17.3%; Score 32; DB 22; Length 139;
 XX Best Local Similarity 100.0%; Pred. No. 5.2e-25;
 XX Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 65 TDGFCRSAPVTELVCSGCGPARLLPNAIGR 96
 Db 16 TDGFCRSAPVTELVCSGCGPARLLPNAIGR 47
 RESULT 6
 ID ABG42911 standard; Peptide; 139 AA.
 AC ABG42911;
 XX
 XX 19-AUG-2002 (first entry)
 XX Human peptide encoded by genome-derived single exon probe SEQ ID 32576.
 XX Human; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 XX Homo sapiens.
 XX WO200186003-A2.
 XX

PD 15-NOV-2001.
 XX 30-JAN-2001; 2001WO-US006665.
 XX 04-FEB-2000; 2000US-180312P.
 XX 26-MAY-2000; 2000US-207456P.
 XX 30-JUN-2000; 2000US-0608408.
 XX 03-AUG-2000; 2000US-0632366.
 XX 21-SEP-2000; 2000US-234687P.
 XX 27-SEP-2000; 2000US-236359P.
 XX 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2002-114183/15.
 XX Spatially-addressable set of single exon nucleic acid probes, used to
 XX measure gene expression in human lung samples -
 XX Claim 27; SEQ ID NO 32576; 634pp; English.
 XX The invention relates to a spatially-addressable set of single exon
 XX nucleic acid probes for measuring gene expression in a sample derived
 XX from human lung comprising single exon nucleic acid probes having one of
 XX 12614 nucleic acid sequences mentioned in the specification, or their
 XX complements or the 1387 open reading frames derived from the 12614
 XX probes. Also included are a microarray comprising the novel set of
 XX probes; the novel set of probes which hybridise at high stringency to a
 XX nucleic acid expressed in the human lung; measuring gene expression in a
 XX sample derived from human lung, comprising (a) contacting the array with
 XX a collection of detectably labeled nucleic acids derived from human lung
 XX mRNA, and (b) measuring the label detectably bound to each probe of
 XX the array; identifying exons in a eukaryotic genome, comprising
 XX (a) algorithmically predicting at least one exon from genomic sequences
 XX of the eukaryote; and (b) detecting specific hybridisation of detectably
 XX labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 XX having a fragment identical to the predicted exon, the probe is included
 XX in the above mentioned microarray; assigning exons to a single gene,
 XX comprising (a) identifying exons from genomic sequence by the method
 XX above and (b) measuring the expression of each of the exons in several
 XX tissues and/or cell types using hybridisation to a single exon
 XX microarrays having a probe with the exon, where a common pattern of
 XX expression of the exons in the tissues and/or cell types indicates that
 XX the exons should be assigned to a single gene; a peptide comprising one
 XX of 12011 sequences, mentioned in the specification, or encoded by the
 XX probes/open reading frames (ORF). The probes are used for gene
 XX expression analysis, and for identifying exons in a gene, particularly
 XX using human lung derived mRNA and for the study of lung diseases
 XX such as asthma, lung cancer, chronic obstructive pulmonary disease
 XX (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 XX fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 XX Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 XX haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 XX pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
 XX pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 XX and hyaline membrane disease. The present sequence is a peptide/protein
 XX encoded by a single exon probe of the invention.
 XX Note: The sequence data for this patent did not form part
 XX of the printed specification, but was obtained in electronic
 XX format directly from WIPO at
 XX ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 139 AA;
 XX

Query Match 17.3%; Score 32; DB 23; Length 139;
 Best Local Similarity 100.0%; Pred. No. 5.2e-25;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 65 TDGFCRSAPVTELVCSGCGPARLLPNAIGR 96
 Db 16 TDGFCRSAPVTELVCSGCGPARLLPNAIGR 47

AC	AA96433;
XX	
TT	12-SEP-2000 (first entry)
DT	
XX	
DE	Rat TGF-beta binding protein (BEER).
XX	
KW	osteopathic; transforming growth factor-beta; TGF-beta; binding protein; BEER; gene therapy; antisense therapy; fracture; bone mineralization.
KK	
WW	
OS	Rattus norvegicus.
XX	
PB	WO200032773-A1.
NN	
PN	08-JUN-2000.
XX	
PD	
DD	24-NOV-1999; 99WO-US27990.
PF	
PP	
XX	
PR	27-NOV-1998; 98US-0110283.
XX	
PA	(DARW-) DARWIN DISCOVERY LTD.
XX	
PI	Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paeper BW;
PI	Van Ness J, Winkler DG;
XX	
XX	WPI: 2000-412321/35.
DR	N-PSTDB; AAA29059.
XX	
DX	
XX	Nucleic acids (I) encoding a transforming growth factor beta binding protein, useful for identifying agents for treating osteopenia, osteoporosis and fractures
PT	
PT	Claim 6; Page 125-126; 162pp: English.
PS	
XX	This shows a rat transforming growth factor-beta (TGF-beta) binding protein designated rBEER. The cDNA and protein may be used for prevention, treatment and diagnosis of diseases associated with CC inappropriate BEER expression. For example, they may be used to treat disorders associated with decreased TGF-beta BP expression. The cDNA or vectors may be administered to treat diseases by rectifying mutations or deletions in a patient's genome that affect the activity of BEER by expressing inactive proteins or to supplement the patients own production of BEER polypeptides. The nucleic acids may be used for recombinant production of BEER, gene therapy, antisense therapy, as probes for diagnostic assays and for functional studies. BEER may be used to raise antibodies and for identification of BEER modulators. BEER antagonists may be used to increase bone mineral content for the treatment of disorders such as osteopenia, osteoporosis, fractures and other disorders associated with low mineral content.
QQ	
SQ	Sequence 213 AA;
Query Match	58.4%; Score 108; DB 21; Length 213;
Best Local Similarity	100.0%; Pred.No. 9e-104;
Mismatches	0; Mismatches 0; Indels 0; Gaps 0;
Matches	108; Conservative 0;
QY	65 TDGPCRSKAPVTVELVCSCGCGPARLLPNAIGRKVKWRRENGDPFCIPDHYRAQRVOLLC ¹²⁴
Db	90 TDGPCRSKAPVTVELVCSCGCGPARLLPNAIGRKVKWRRENGDFRCIPDHYRAQRVOLLC ¹⁴⁹
OY	125 GGAAPRSRRVLIVASCKKRLTRFHNQSELKDFGPETARPQGKRKRP ¹⁷²
Db	150 GGAAPRSRRVLIVASCKKRLTRFHNQSELKDFGPETARPQGKRKRP ¹⁹⁷
RESULT 5	
AM73067	
ID	AA73067 standard; Protein; 139 AA.
XX	
AC	AA73067;
XX	
DT	06-NOV-2001 (first entry)
DF	Human bone marrow expressed probe encoded protein SEQ ID NO: 33373.

PA (AMGE-) AMGEN INC.
 PI Paszty CJ, Gao Y;
 XX
 XX
 DR WPI; 2002-114325/15.
 DR N-PSDB; ABA94294.
 XX
 XX New human and mouse cysteine-knot polypeptide designated as Cloaked-2,
 PT for treating or preventing kidney, heart (e.g. myocardial infarction)
 PT or liver (e.g. hepatitis) diseases
 XX
 XX Claim 13; Fig 2; 170pp; English.
 XX
 CC The invention relates to polypeptides comprising a cysteine knot motif
 CC and designated as Cloaked-2, derived from human and mouse. The cloaked-2
 CC polypeptides can be expressed by standard recombinant methodology. The
 CC cloaked-2 polynucleotides are useful in gene therapy and antisense
 CC therapy. The cloaked-2 polypeptides and polynucleotides are useful for
 CC treating, preventing, ameliorating or detecting diseases and disorders of
 CC the kidney (e.g. anemia, hypertension or low blood pressure), heart (e.g.
 CC cardiac hypertrophy, congestive heart failure, myocardial infarction,
 CC arrhythmias, atherosclerosis, hypertension or low blood pressure),
 CC skeletal muscle (e.g. muscular dystrophy or cachexia), placenta (e.g.
 CC congenital abnormalities or miscarriage), liver (e.g. hepatitis or
 CC cirrhosis), pancreas (e.g. diabetes or pancreatitis), thyroid (e.g.
 CC Grave's disease or myxedema) or adrenal cortex (e.g. Cushing's disease
 CC cancer or myopathies), infections, or autoimmune diseases. Selective
 CC binding agents may be used to modulate the biological activities of
 CC Cloaked-2 polypeptides or to detect Cloaked-2 polypeptide levels in a
 CC sample. Transgenic non-human animals are useful for drug candidate
 CC screening. The present sequence represents the mouse cloaked-2 mature
 CC polypeptide.
 XX
 SQ Sequence 188 AA;
 Query Match 94.1%; Score 174; DB 23; Length 188;
 Best Local Similarity 100.0%; Pred. No. 2.4e-172;
 Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QGWAQFRNDATEVTPGLGEYPEPPENNQTMRAENGRRPHHPYDAKDVSSEYSCRELHY 60
 Db 1 QGWAQFRNDATEVTPGLGEYPEPPENNQTMRAENGRRPHHPYDAKDVSSEYSCRELHY 60
 QY 61 TRFLTDPGCRSAKPVTELVCSCGCGPARLLPNAIGRVKWRPNPGRPFRCIPDRYRAQRVQ 120
 Db 61 TRFLTDPGCRSAKPVTELVCSCGCGPARLLPNAIGRVKWRPNPGRPFRCIPDRYRAQRVQ 120
 QY 121 LLCPGAAPRSRKVRVLVASCCKKRLTRFHNSQLKDFGPETARPOKGRKPRPGA 174
 Db 121 LLCPGAAPRSRKVRVLVASCCKKRLTRFHNSQLKDFGPETARPOKGRKPRPGA 174
 RESULT 2
 AAY96432
 ID AAY96432 standard; Protein; 211 AA.
 XX
 XX AAY96432;
 XX
 DT 12-SEP-2000 (first entry)
 XX
 DE Murine TGF-beta binding protein (BEER).
 XX
 KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KW BEER; gene therapy; antisense therapy; fracture; bone mineralization.
 XX
 OS Mus musculus.
 XX
 PN WO200032773-A1.
 XX
 XX 08-JUN-2000.
 PD
 XX 24-NOV-1999; 99WO-US27990.
 PF

XX 27-NOV-1998; 98US-0110283.
 XX (DAEW-) DARWIN DISCOVERY LTD.
 XX
 XX Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepers BW;
 PI Van Ness J, Winkler DG;
 XX WPI; 2000-412321/35.
 DR N-PSDB; AAA29058.
 XX
 XX Nucleic acids (1) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures
 XX
 PS Claim 5; Page 124; 152pp; English.
 XX
 CC This shows a murine transforming growth factor-beta (TGF-beta)
 CC binding protein designated mBEER. The cDNA and protein may be used for
 CC prevention, treatment and diagnosis of diseases associated with
 CC inappropriate BEER expression. For example, they may be used to treat
 CC disorders associated with decreased TGF-beta BP expression. The cDNA or
 CC vectors may be administered to treat diseases by rectifying mutations or
 CC deletions in a patient's genome that affect the activity of BEER by
 CC expressing inactive proteins or to supplement the patients own production
 CC of BEER polypeptides. The nucleic acids may be used for recombinant
 CC production of BEER, gene therapy, antisense therapy, as probes for
 CC diagnostic assays and for functional studies. BEER may be used to raise
 CC antibodies and for identification of BEER modulators. BEER antagonists
 CC may be used to increase bone mineral content for the treatment of
 CC disorders such as osteopenia, osteoporosis, fractures and other disorders
 CC associated with low mineral content.
 XX
 SQ Sequence 211 AA;
 Query Match 94.1%; Score 174; DB 21; Length 211;
 Best Local Similarity 100.0%; Pred. No. 2.6e-172;
 Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QGWAQFRNDATEVTPGLGEYPEPPENNQTMRAENGRRPHHPYDAKDVSSEYSCRELHY 60
 Db 24 QGWAQFRNDATEVTPGLGEYPEPPENNQTMRAENGRRPHHPYDAKDVSSEYSCRELHY 83
 QY 61 TRFLTDPGCRSAKPVTELVCSCGCGPARLLPNAIGRVKWRPNPGRPFRCIPDRYRAQRVQ 120
 Db 84 TRFLTDPGCRSAKPVTELVCSCGCGPARLLPNAIGRVKWRPNPGRPFRCIPDRYRAQRVQ 143
 QY 121 LLCPGAAPRSRKVRVLVASCCKKRLTRFHNSQLKDFGPETARPOKGRKPRPGA 174
 Db 144 LLCPGAAPRSRKVRVLVASCCKKRLTRFHNSQLKDFGPETARPOKGRKPRPGA 197
 RESULT 3
 ABB07210
 ID ABB07210 standard; Protein; 211 AA.
 XX
 XX ABB07210;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Mouse cloaked-2 polypeptide sequence.
 XX
 KW Cloaked-2; cysteine knot motif; nephrotropic; cardiant; immunomodulator;
 KW hepatotropic; antiinflammatory; antithyroid; cytostatic; neuroprotective;
 KW antianemic; hypotensive; antiarrhythmic; antiarteriosclerotic; muscular;
 KW antidiabetic; anorectic; gene therapy; cell therapy; antisense therapy;
 KW mouse.
 XX
 OS Mus musculus.
 XX
 XX Key Location/Qualifiers
 FT Peptide 1..23
 FT /note= "signal peptide"

Result No.	Score	% Match	Length	DB	ID	Description
1	174	94.1	188	23	AB807208	Mouse cloaked-2 po
2	174	94.1	211	21	AAV96432	Murine TGF-beta bi
3	174	94.1	211	23	AB807210	Mouse cloaked-2 po
4	108	58.4	213	21	AAV96433	Rat TGF-beta bindi
5	32	17.3	139	22	AAW73067	Human bone marrow
6	32	17.3	139	23	ABG42911	Human peptide enco
7	32	17.3	176	21	AAV96434	Bovine TGF-beta bi
8	32	17.3	190	23	AB807207	Human cloaked-2 po
9	32	17.3	213	21	AAE26106	Human DAN/Cerberus
10	32	17.3	213	21	AAV96439	Human TGF-beta bin

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	174	94.1	188	23	AB807208	Mouse cloaked-2 po
2	174	94.1	211	21	AAV96432	Murine TGF-beta bi
3	174	94.1	211	23	AB807210	Mouse cloaked-2 po
4	108	58.4	213	21	AAV96433	Rat TGF-beta bindi
5	32	17.3	139	22	AAW73067	Human bone marrow
6	32	17.3	139	23	ABG42911	Human peptide enco
7	32	17.3	176	21	AAV96434	Bovine TGF-beta bi
8	32	17.3	190	23	AB807207	Human cloaked-2 po
9	32	17.3	213	21	AA826106	Human DAN/Cerberus
10	32	17.3	213	21	AAV96429	Human TGF-beta bin

PR 01-JUN-2000; 2000US-223542P.
PR 04-AUG-2000; 2000US-223542P.

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	174	94.1	188	23	AB807208	Mouse cloaked-2 po
2	174	94.1	211	21	AAV96432	Murine TGF-beta bi
3	174	94.1	211	23	AB807210	Mouse cloaked-2 po
4	108	58.4	213	21	AAV96433	Rat TGF-beta bindi
5	32	17.3	139	22	AAW73067	Human bone marrow
6	32	17.3	139	23	ABG42911	Human peptide enco
7	32	17.3	176	21	AAV96434	Bovine TGF-beta bi
8	32	17.3	190	23	AB807207	Human cloaked-2 po
9	32	17.3	213	21	AA826106	Human DAN/Cerberus
10	32	17.3	213	21	AAV96429	Human TGF-beta bin

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/10/159,749
FILING DATE: 29-May-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/040,229
FILING DATE: 13-Mar-1998
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B97-020-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 184 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-159-749-6

Query Match 8.8%; Score 90.5; DB 9; Length 184;
Best Local Similarity 22.9%; Pred. No. 0.21; Mismatches 60; Indels 35; Gaps 8;
Matches 35; Conservative 23;

QY 18 GEYEP-----PPENNTMNAENGSRPHH-----PYDAKDVSEYSCRELHYT--RFLTDG 67
Db 33 GAIPPPDKQPNDSQMQTQQQSGHRERGKGTSMPEAEVLESSQEAHLHITERKYLKED 92
QY 68 PCR-----SAKPVTELCSGCGPARLLPNAIGRVKWRPNPDPFR-----CI 110
Db 93 WCKTQPLKQTIHEEGCNSRTIINRFYGCNC-SFYIPRHV-----RKEEGSFQSCSFC 145
QY 111 PDYRAQRVQLLCPGGAAPRSRK-VRLVASCKC 142
Db 146 PKKFTTMTLNCPELQPPKRRKRIITRVKRC 178

RESULT 14
US-10-159-749-4
Sequence 4, Application US/10159749
Publication No. US20020192219A1
GENERAL INFORMATION:
APPLICANT: Harland, Richard
Hsu, David
TITLE OF INVENTION: Morphogenic Proteins
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/159,749
FILING DATE: 29-May-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/040,229
FILING DATE: 13-Mar-1998
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: B97-020-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 182 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-159-749-4

Query Match 8.6%; Score 88.5; DB 9; Length 182;
Best Local Similarity 23.5%; Pred. No. 0.32; Mismatches 54; Indels 37; Gaps 8;
Matches 35; Conservative 23;

QY 23 PPENNTMNAENG-----GRPPHPYDAKDVSEYSCRELHYT--RFLTDGPCR- 70
Db 36 PPPDKGP-NDSEOGCAQPGDRVRGKGQALAAEEVLESSQEAHLHITERKYLKRDWCKT 94
QY 71 -----SAKPVTELCSGCGPARLLPNAIGRVKWRPNPDPFR-----CIPDRY 114
Db 95 QPLKQTIHEEGCNSRTIINRFYGCNC-SFYIPRHIRE-----GSFQSCSFCCKPKF 147
QY 115 RAQRVQLLCPGGAAP-RSRKRVLVASCKC 142
Db 148 TTMVTLNCPQLPPTKKRRIITRVKQRC 176

RESULT 15
US-10-044-716-10
Sequence 10, Application US/10044716
Patent No. US20020159986A1
GENERAL INFORMATION:
APPLICANT: LANGENFELD, John
TITLE OF INVENTION: BONE MORPHOGENETIC PROTEIN-2 IN THE TREATMENT AND DIAGNOSIS OF CANCER
FILE REFERENCE: 270/070US
CURRENT APPLICATION NUMBER: US/10/044,716
CURRENT FILING DATE: 2002-01-11
PRIOR APPLICATION NUMBER: US60/261,252
PRIOR FILING DATE: 2001-01-12
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Patent in version 3.1
SEQ ID NO 10
LENGTH: 184
TYPE: PRT
ORGANISM: Homo sapiens
US-10-044-716-10

Query Match 8.6%; Score 88.5; DB 9; Length 184;
Best Local Similarity 23.4%; Pred. No. 0.32; Mismatches 55; Indels 43; Gaps 8;
Matches 36; Conservative 20;

QY 23 PPENNTMNAENG-----TMNAENGCRPHHPYDAKDVSEYSCRELHYT--RFLTD 66
Db 34 PPPDKAHNDSEQTQSPQPGSRNRGRGCGRTAMP--GEEVLESSQEAHLHITERKYLK 91
QY 67 GPCR-----SAKPVTELCSGCGPARLLPNAIGRVKWRPNPDPFR-----C 109
Db 92 DWCKTQPLKQTIHEEGCNSRTIINRFYGCNC-SFYIPRHIRE-----RKEEGSFQSCSFC 144
QY 110 IPDYRAQRVQLLCPGGAAP-RSRKRVLVASCKC 142
Db 145 PKKFTTMTLNCPELQPPKRRKRIITRVKQRC 178

Search completed: March 28, 2003, 14:20:45
Job time : 11.84 secs

US-09-867-274-25

; Sequence 25, Application US/0986774

; Patent No. US20020106650A1

; GENERAL INFORMATION:

; APPLICANT: Paszty, Christopher

; APPLICANT: Gao, Yongming

; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof

; FILE REFERENCE: 01017/37428

; CURRENT APPLICATION NUMBER: US/09/867,274

; CURRENT FILING DATE: 2001-05-29

; PRIOR APPLICATION NUMBER: US 60/208,550

; PRIOR FILING DATE: 2000-06-01

; PRIOR APPLICATION NUMBER: US 60/223,542

; PRIOR FILING DATE: 2000-08-04

; NUMBER OF SEQ ID NOS: 25

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 25

; LENGTH: 183

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-867-274-25

Query Match

Best Local Similarity 31.4%; Score 324; DB 10; Length 183;

Matches 72; Conservative 36; Mismatches 62; Indels 12; Gaps 7;

QY 6 FRNDATVIGLGEYPPPP-PENNQTWNRAENGGRPHHPYDAKDVDS---EYSCRELHYT 61

Db 1 FKNDATVILYSHVVKVPAHPSSNSTLNQANGR--HFSNTGLDNRTRVQVGCRLRST 58

QY 62 RFLTDGCRSAKPVTELVCSCQCQGPALLPNAIG--RVKWW-RPNGPDPFCIPDRYRAQ 117

Db 59 KYISDGQCTSIPLKELVACGECPLPVLPNWIGGYGTKYWSRSQEWKVCNDKTRTQ 118

QY 118 RVQLLPCGGAAPRRKRVLVASCKKRLTRFHNSQLKDFGPETARP-QKGRKRPQAKA 176

Db 119 RIQLQCDGST-RYKITYVTWACKCKRYTROHNSHNFESMSPAKVPVQHHRKXASKS 177

QY 177 NQ 178

Db 178 SK 179

RESULT 11

US-09-814-777A-126

; Sequence 126, Application US/09814777A

; Patent No. US20020142415A1

; GENERAL INFORMATION:

; APPLICANT: KOOPMAN, Peter Anthony

; APPLICANT: MUSCAT, George Eugene Orlando

; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND POLYNUCLEOTIDES AND METHODS OF USING THEM

; FILE REFERENCE: 21415-0003

; CURRENT APPLICATION NUMBER: US/09/814,777A

; CURRENT FILING DATE: 2001-03-23

; PRIOR APPLICATION NUMBER: AU PQ6457

; PRIOR FILING DATE: 2000-03-24

; NUMBER OF SEQ ID NOS: 128

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 126

; LENGTH: 337

; TYPE: PRT

; ORGANISM: Human

US-09-814-777A-126

Query Match

Best Local Similarity 10.7%; Score 110; DB 10; Length 337;

Matches 50; Conservative 21; Mismatches 61; Indels 64; Gaps 10;

QY 1 QGWAQFRNDATVIGLGEYPPPPENNQTWNRAENGGRPHHPYDAKDVSEYSCRELHY 60

Db 117 QARKARLEPGLLLPGLAP-PQPPPE-----PFPASGASARAFRELPP 158

QY 61 TRFLTGDG---PCRSAPVTELVCSCQCQGPALL-PNAIGRVKWWRPNGDPFCIPDRYRA 116

Db 159 LGAEFDGLGLTPERSPLDGL-----EPGEAAFPFPAPRTARWRPSAP-----PTAHRV 209
QY 117 ---QVQLLCPGGA---APRSRKRVLVAS-----CKKRLTRFHNSQLKDFGP 159
Db 210 VAGPRLLRSGSGGAQDRAPRARSACTTAPWARPARTPARCRR----- 254
QY 160 ETAPQKGRKRPQAK 175
Db 255 -RPRRRWRAPSPGAR 269

RESULT 12

US-09-814-777A-99

; Sequence 99, Application US/09814777A

; Patent No. US20020142415A1

; GENERAL INFORMATION:

; APPLICANT: KOOPMAN, Peter Anthony

; APPLICANT: MUSCAT, George Eugene Orlando

; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND POLYNUCLEOTIDES AND METHODS OF USING THEM

; FILE REFERENCE: 21415-0003

; CURRENT APPLICATION NUMBER: US/09/814,777A

; CURRENT FILING DATE: 2001-03-23

; PRIOR APPLICATION NUMBER: AU PQ6457

; PRIOR FILING DATE: 2000-03-24

; NUMBER OF SEQ ID NOS: 128

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 99

; LENGTH: 462

; TYPE: PRT

; ORGANISM: Human

US-09-814-777A-99

Query Match

Best Local Similarity 9.3%; Score 95.5; DB 10; Length 462;

Matches 45; Conservative 23; Mismatches 64; Indels 63; Gaps 10;

QY 1 QGWAQFRNDATVIGLGEYPPPPENNQTWNRAENG--GRPPHPYDAKDVSEYSCREL 58

Db 117 QARKARLEPGLLLPGLAP-PQPPPEPFPASGASARAFSGPAGAEFDGLGL----- 167

QY 59 HVTRFLTDGCRSAKPVTELVCSCQCQGPALL-PNAIGRVKWWRPNGDPFCIPDRYRA- 116

Db 168 -----PTPERSPLDGL-----EPGEAAFPFPAPRTARWRPSAP-----PTAHRV 209

QY 117 ---QVQLLCPGGA---APRSRKRVLVAS-----CKKRLTRFHNSQLKDFGP 160

Db 210 VAGPRLLRSGSGGAQDRAPRARSACTTAPWARPARTPARCRR----- 253

QY 161 TAPQKGRKRPQAK 175

Db 254 RPRRRWRAPSPGAR 268

RESULT 13

US-10-159-749-6

; Sequence 6, Application US/10159749

; Publication No. US20020192219A1

; GENERAL INFORMATION:

; APPLICANT: Harland, Richard

; APPLICANT: Hsu, David

; TITLE OF INVENTION: Morphogenic Proteins

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

; STREET: 75 DENISE DRIVE

; CITY: HILLSBOROUGH

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94010

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

QY 61 TRFLTDGCRSAKPVTELVCSCGCGPARLLPNAIG---RVKWW-RPNGDPDFRCIPDRYRA 116
Db 81 TKYISDQCTSIISPLKELVCAGECLPLPLPNWIGGGYGTKYWSRRSQEWRCVNDKTRT 140
QY 117 ORVOLLCPGGAAPRRKVLVASCCKKLTFRHNSQELKDFGPETARP-QKGRKPRPGAK 175
Db 141 ORIQLOQODGST-RTYKITVVTACKRYTROHNESSHNFESMSPAKPVQHHRRKRAK 199
QY 176 ANQAL 181
Db 200 SSKSL 205
RESULT 8
US-09-853-625B-2
; Sequence 2, Application US/09853625B
; Patent No. US20020049304A1
; GENERAL INFORMATION:
; APPLICANT: Hastings, Gregg A. and Adams, Mark D.
; TITLE OF INVENTION: Human CCN-Like Growth Factor
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/853,625B
; FILING DATE: 14-May-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/053,587
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-442
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 206 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS: <Unknown>
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-853-625B-2
Query Match 31.8%; Score 328; DB 10; Length 206;
Best Local Similarity 39.9%; Pred. No. 2.5e-24;
Matches 73; Conservative 36; Mismatches 62; Indels 12; Gaps 7;
QY 5 AFNRDATEVPIGLGEYPEPP-PENNOTMRAENGRRPHHPYDAKDVS---EYSCRELHY 60
Db 23 AFKNDATEILYSHVVKPVPAHSSNSTLNQANGRR--HFSNTGLDRNTRVQVGCRLRS 80
QY 61 TRFLTDGCRSAKPVTELVCSCGCGPARLLPNAIG---RVKWW-RPNGDPDFRCIPDRYRA 116
Db 81 TKYISDQCTSIISPLKELVCAGECLPLPLPNWIGGGYGTKYWSRRSQEWRCVNDKTRT 140
QY 117 ORVOLLCPGGAAPRRKVLVASCCKKLTFRHNSQELKDFGPETARP-QKGRKPRPGAK 175
Db 141 ORIQLOQODGST-RTYKITVVTACKRYTROHNESSHNFESMSPAKPVQHHRRKRAK 199

QY 176 ANQ 178
Db 200 SSK 202
RESULT 9
US-09-853-625B-20
; Sequence 20, Application US/09853625B
; Patent No. US20020049304A1
; GENERAL INFORMATION:
; APPLICANT: Hastings, Gregg A. and Adams, Mark D.
; TITLE OF INVENTION: Human CCN-Like Growth Factor
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/853,625B
; FILING DATE: 14-May-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/053,587
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-442
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 206 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS: <Unknown>
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-853-625B-20
Query Match 31.8%; Score 328; DB 10; Length 206;
Best Local Similarity 39.9%; Pred. No. 2.5e-24;
Matches 73; Conservative 36; Mismatches 62; Indels 12; Gaps 7;
QY 5 AFNRDATEVPIGLGEYPEPP-PENNOTMRAENGRRPHHPYDAKDVS---EYSCRELHY 60
Db 23 AFKNDATEILYSHVVKPVPAHSSNSTLNQANGRR--HFSNTGLDRNTRVQVGCRLRS 80
QY 61 TRFLTDGCRSAKPVTELVCSCGCGPARLLPNAIG---RVKWW-RPNGDPDFRCIPDRYRA 116
Db 81 TKYISDQCTSIISPLKELVCAGECLPLPLPNWIGGGYGTKYWSRRSQEWRCVNDKTRT 140
QY 117 ORVOLLCPGGAAPRRKVLVASCCKKLTFRHNSQELKDFGPETARP-QKGRKPRPGAK 175
Db 141 ORIQLOQODGST-RTYKITVVTACKRYTROHNESSHNFESMSPAKPVQHHRRKRAK 199
QY 176 ANQ 178
Db 200 SSK 202
RESULT 10

RESULT 2

US-09-867-274-6

; Sequence 6, Application US/09867274

; Patent No. US20020106650A1

; GENERAL INFORMATION:

; APPLICANT: Paszty, Christopher

; APPLICANT: Gao, Yongming

; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof

; FILE REFERENCE: 01017/37428

; CURRENT APPLICATION NUMBER: US/09/867,274

; CURRENT FILING DATE: 2001-05-29

; PRIOR APPLICATION NUMBER: US 60/208,550

; PRIOR FILING DATE: 2000-06-01

; PRIOR APPLICATION NUMBER: US 60/223,542

; PRIOR FILING DATE: 2000-08-04

; NUMBER OF SEQ ID NOS: 25

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 6

; LENGTH: 208

; TYPE: PRT

; ORGANISM: Mus musculus

US-09-867-274-6

Query Match

Best Local Similarity 100.0%; Score 1032; DB 10; Length 208;

Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

1 QGWAQFRNDATFVIFGLGEYEPPEPP--ENNQTWNAENGRRPHHPYDAKDVSEYSCRELHY 60

Db

24 QGWAQFRNDATFVIFGLGEYEPPEPP--ENNQTWNAENGRRPHHPYDAKDVSEYSCRELHY 83

QY

61 TRFLTDGFCRSAPKPVTELVCSCGCGPARLLPNAIGRVKWRPNPGRPCIPDRYRAQRVQ 120

Db

84 TRFLTDGFCRSAPKPVTELVCSCGCGPARLLPNAIGRVKWRPNPGRPCIPDRYRAQRVQ 143

QY

121 LLCPGGAAPRRKRVLVASCKCKELTRFHNSQELKDFGTEAARPKQGRKPRPARSAK 180

Db

144 LLCPGGAAPRRKRVLVASCKCKELTRFHNSQELKDFGTEAARPKQGRKPRPARSAK 203

QY

181 LENAY 185

Db

204 LENAY 208

RESULT 3

US-09-867-274-2

; Sequence 2, Application US/09867274

; Patent No. US20020106650A1

; GENERAL INFORMATION:

; APPLICANT: Paszty, Christopher

; APPLICANT: Gao, Yongming

; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof

; FILE REFERENCE: 01017/37428

; CURRENT APPLICATION NUMBER: US/09/867,274

; CURRENT FILING DATE: 2001-05-29

; PRIOR APPLICATION NUMBER: US 60/208,550

; PRIOR FILING DATE: 2000-06-01

; PRIOR APPLICATION NUMBER: US 60/223,542

; PRIOR FILING DATE: 2000-08-04

; NUMBER OF SEQ ID NOS: 25

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 2

; LENGTH: 190

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-867-274-2

Query Match

Best Local Similarity 90.4%; Score 932.5; DB 10; Length 190;

Matches 168; Conservative 10; Mismatches 7; Indels 5; Gaps 2;

QY 1 QGWAQFRNDATFVIFGLGEYEPPEPP--ENNQTWNAENGRRPHHPYDAKDVSEYSCREL 58

Db 1 QGWAQFRNDATFVIFGLGEYEPPEPP--ENNQTWNAENGRRPHHPYDAKDVSEYSCREL 60

QY 59 HYTRFLTDGFCRSAPKPVTELVCSCGCGPARLLPNAIGRVKWRPNPGRPCIPDRYRAQR 118

Db 61 HFTRYVTDGFCRSAPKPVTELVCSCGCGPARLLPNAIGRVKWRPNPGRPCIPDRYRAQR 120

QY 119 VOLLCPGGAAPRRKRVLVASCKCKELTRFHNSQELKDFGTEAARPKQGRKPRPARSAK 175

Db 121 VOLLCPGGAAPRRKRVLVASCKCKELTRFHNSQELKDFGTEAARPKQGRKPRPARSAK 180

QY 176 ANQAELENAY 185

Db 181 ANQAELENAY 190

RESULT 4

US-09-867-274-5

; Sequence 5, Application US/09867274

; Patent No. US20020106650A1

; GENERAL INFORMATION:

; APPLICANT: Paszty, Christopher

; APPLICANT: Gao, Yongming

; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof

; FILE REFERENCE: 01017/37428

; CURRENT APPLICATION NUMBER: US/09/867,274

; CURRENT FILING DATE: 2001-05-29

; PRIOR APPLICATION NUMBER: US 60/208,550

; PRIOR FILING DATE: 2000-06-01

; PRIOR APPLICATION NUMBER: US 60/223,542

; PRIOR FILING DATE: 2000-08-04

; NUMBER OF SEQ ID NOS: 25

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 5

; LENGTH: 213

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-867-274-5

Query Match

Best Local Similarity 90.4%; Score 932.5; DB 10; Length 213;

Matches 168; Conservative 10; Mismatches 7; Indels 5; Gaps 2;

QY

1 QGWAQFRNDATFVIFGLGEYEPPEPP--ENNQTWNAENGRRPHHPYDAKDVSEYSCREL 58

Db

24 QGWAQFRNDATFVIFGLGEYEPPEPP--ENNQTWNAENGRRPHHPYDAKDVSEYSCREL 83

QY

59 HYTRFLTDGFCRSAPKPVTELVCSCGCGPARLLPNAIGRVKWRPNPGRPCIPDRYRAQR 118

Db

84 HFTRYVTDGFCRSAPKPVTELVCSCGCGPARLLPNAIGRVKWRPNPGRPCIPDRYRAQR 143

QY

119 VOLLCPGGAAPRRKRVLVASCKCKELTRFHNSQELKDFGTEAARPKQGRKPRPARSAK 175

Db

144 VOLLCPGGAAPRRKRVLVASCKCKELTRFHNSQELKDFGTEAARPKQGRKPRPARSAK 203

QY

176 ANQAELENAY 185

Db

204 ANQAELENAY 213

RESULT 5

US-09-864-761-47109

; Sequence 47109, Application US/09864761

; Patent No. US20020048763A1

; GENERAL INFORMATION:

; APPLICANT: Penn, Sharon G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; APPLICANT: Chen, Wenheng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

; FILE REFERENCE: Aemica-X-1

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:16:22 ; Search time 11.84 Seconds
(without alignments)
917.557 Million cell updates/sec

Title: US-09-867-274-4
Perfect score: 1032
Sequence: 1 QGQAFRNDATEVIFGLGEY.....KGRKPRCAKANQALENAY 185

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 237916 seqs, 58723674 residues

Total number of hits satisfying chosen parameters: 237916

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1032	100.0	185	10	US-09-867-274-4
2	1032	100.0	208	10	US-09-867-274-6
3	932.5	90.4	190	10	US-09-867-274-2
4	932.5	90.4	213	10	US-09-867-274-5
5	687.5	66.6	139	10	US-09-864-761-47109
6	330	32.0	206	9	US-09-866-050A-159
7	330	32.0	206	9	US-09-866-050A-286
8	328	31.8	206	10	US-09-853-625B-2
9	328	31.8	206	10	US-09-853-625B-20
10	324	31.4	183	10	US-09-867-274-25
11	110	10.7	337	10	US-09-814-777A-126
12	95.5	9.3	462	10	US-09-814-777A-99
13	90.5	8.8	184	9	US-10-159-749-6
14	88.5	8.6	182	9	US-10-159-749-4
15	88.5	8.6	184	9	US-10-044-716-10
16	88.5	8.6	184	9	US-10-159-749-2
17	86	8.3	187	9	US-10-159-749-9
18	83.5	8.1	147	9	US-10-092-154-828
19	83.5	8.1	147	10	US-09-764-847-828

20	80.5	7.8	270	9	US-09-903-170C-1
21	80.5	7.8	270	10	US-09-903-180B-1
22	80.5	7.8	270	10	US-09-903-187A-1
23	80.5	7.8	270	10	US-09-903-171A-1
24	80.5	7.8	270	10	US-09-903-188A-1
25	80.5	7.8	270	10	US-09-903-323A-1
26	80.5	7.8	270	10	US-09-903-325A-1
27	80	7.8	188	10	US-09-349-954A-18
28	80	7.8	188	10	US-09-907-007-18
29	80	7.8	251	10	US-09-764-853-522
30	80	7.8	251	10	US-09-764-898-176
31	76.5	7.4	141	9	US-09-813-398-3
32	76.5	7.4	145	9	US-09-760-294-12
33	76.5	7.4	145	9	US-09-760-294-13
34	76.5	7.4	145	9	US-09-760-294-14
35	76.5	7.4	145	9	US-09-760-294-15
36	76.5	7.4	145	9	US-10-131-241-65
37	76.5	7.4	165	9	US-09-915-676-1
38	76.5	7.4	165	10	US-09-466-320-14
39	76.5	7.4	267	9	US-10-044-716-12
40	76.5	7.4	267	9	US-09-089-818B-8
41	76.5	7.4	270	9	US-10-072-349-145
42	76.5	7.4	270	10	US-09-764-855-145
43	76.5	7.4	307	10	US-09-756-186-4
44	76.5	7.4	336	10	US-09-756-186-8
45	76.5	7.4	432	12	US-10-062-254-242

ALIGNMENTS

RESULT 1
US-09-867-274-4
; Sequence 4, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; APPLICANT: Gao, Yongming
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-867-274-4

Query Match	100.0%	Score	1032;	DB	10;	Length	185;
Best Local Similarity	100.0%	Pred. No.	1.7e-92;				
Matches	185;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
QY	1	QGWQAFRNDATEVIFGLGEYPPPPENNQTWNAENGRRPHHPYDAKDYSEYSCRELHY	60				
Db	1	QGWQAFRNDATEVIFGLGEYPPPPENNQTWNAENGRRPHHPYDAKDYSEYSCRELHY	60				
QY	61	TRFLTDGPRCAKPVTELVCSGCGPARLLPNAIGRVKWRPNKGPDPFCIPDRYRAQRVQ	120				
Db	61	TRFLTDGPRCAKPVTELVCSGCGPARLLPNAIGRVKWRPNKGPDPFCIPDRYRAQRVQ	120				
QY	121	LLCPGGAAPRSRKVLVASCCKKGLTRFHNQSELKDFGPETARPOKGRKPRPGAKANQAE	180				
Db	121	LLCPGGAAPRSRKVLVASCCKKGLTRFHNQSELKDFGPETARPOKGRKPRPGAKANQAE	180				
QY	181	LENAY 185					
Db	181	LENAY 185					

; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 3814
; LENGTH: 198285
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 X87344
US-09-880-107-3814

Query Match 3.0%; Score 19; DB 10; Length 198285;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 AAGCCTTCAGGAATGTC 98
|||||
DB 70568 AAGCCTTCAGGAATGTC 70586

RESULT 15

US-09-867-550-675
; Sequence 675, Application US/09867550
; Patent No. US20020082206A1
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Mehraban, Fuad,
; APPLICANT: Conley, Pamela
; APPLICANT: Law, Debbie
; APPLICANT: Topper, James
; TITLE OF INVENTION: No. US20020082206A1 Polynucleotides from Atherogenic Cells and
; FILE REFERENCE: 21402-013 (Cura-313)
; CURRENT APPLICATION NUMBER: US/09/867,550
; CURRENT FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: USSN 60/208,427
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 2125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 675
; LENGTH: 323
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-867-550-675

Query Match 2.8%; Score 18; DB 10; Length 323;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 545 CGGAGACCGCGCGCGC 562
|||||
DB 146 CGGAGACCGCGCGCGC 163

Search completed: March 29, 2003, 10:50:10
Job time : 115.594 secs

Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; APPLICANT: Gao, Yongming
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 19
; LENGTH: 45
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial: PCR primer
US-09-867-274-19

Query Match 3.9%; Score 25; DB 10; Length 45;
Best Local Similarity 100.0%; Pred. No. 0.002;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCAGCCCTCACTAGCCCGTGCC 25
|||||
Db 21 ATGCAGCCCTCACTAGCCCGTGCC 45
|||||

RESULT 8
US-09-867-274-12/c
; Sequence 12, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; APPLICANT: Gao, Yongming
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 24
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial: PCR primer
US-09-867-274-12

Query Match 3.8%; Score 24; DB 10; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.0072;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 241 CTGCACACCCGCTTCTCTGACA 264
|||||
Db 24 CTGCACACCCGCTTCTCTGACA 1
|||||

RESULT 9
US-09-867-274-15
; Sequence 15, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; APPLICANT: Gao, Yongming
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof

FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial: PCR primer
US-09-867-274-15

Query Match 3.3%; Score 21; DB 10; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 247 TACACCCGCTTCTCTGACAGAC 267
|||||
Db 1 TACACCCGCTTCTCTGACAGAC 21
|||||

RESULT 10
US-09-867-274-17
; Sequence 17, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; APPLICANT: Gao, Yongming
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial: PCR primer
US-09-867-274-17

Query Match 3.3%; Score 21; DB 10; Length 21;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 291 GGTCACCGAGTTGGTGCTC 311
|||||
Db 1 GGTCACCGAGTTGGTGCTC 21
|||||

RESULT 11
US-09-867-274-9
; Sequence 9, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; APPLICANT: Gao, Yongming
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01

Qy 535 GACTTCGGG 543
Db 577 GACTTCGGG 585

RESULT 7
US-09-867-274-19
: Sequence 19, Application US/09867274

QY	241	CTGCACTACACCCCGTTCCTGACAGACGGGCCCATGCGCAGCGCCAAAGCGGGTCACCGAG	300
Db	241	CTGCACTACACCCCGTTCCTGACAGACGGGCCCATGCGCAGCGCCAAAGCGGGTCACCGAG	300
QY	301	TTGGTGTGCTCCGGCCAGTCGGCCCCGCGCGGCTGTGTCGCCAACGCCCATCGGGCGCGTG	360
Db	301	TTGGTGTGCTCCGGCCAGTCGGCCCCGCGCGGCTGTGTCGCCAACGCCCATCGGGCGCGTG	360
QY	361	AAGTGTGTGGGCCCGAAGCGAACCGGATTTCGGCTGCATCCCGGATCGCTACCGCGCGGAG	420
Db	361	AAGTGTGTGGGCCCGAAGCGAACCGGATTTCGGCTGCATCCCGGATCGCTACCGCGCGGAG	420
QY	421	CGGGTGCAGCTGTGTCGCCCGCGGGGCGCGCGCGCTCGCGCAAGGTGCGTCTGTGGTG	480
Db	421	CGGGTGCAGCTGTGTCGCCCGCGGGGCGCGCGCGCTCGCGCAAGGTGCGTCTGTGGTG	480
QY	481	GCCTCGTGCAGTGCAGCGCCTCACC CGCTTCCACAACCAAGTCGGAGCTCAAGAGACTTC	540
Db	481	GCCTCGTGCAGTGCAGCGCCTCACC CGCTTCCACAACCAAGTCGGAGCTCAAGAGACTTC	540
QY	541	GGGCGGAGACCGCGCGGCGCGAGAGGGTTCGAAGCGCGGGCCCGCGCGCCCGGGGAGCC	600
Db	541	GGGCGGAGACCGCGCGGCGCGAGAGGGTTCGAAGCGCGGGCCCGCGCGCCCGGGGAGCC	600
QY	601	AAAGCCAAACCAAGCGGAGCTGGAGAACCGCTACTTAG	636
Db	601	AAAGCCAAACCAAGCGGAGCTGGAGAACCGCTACTTAG	636

RESULT 2
US-09-864-761-14440/c
Sequence 14440, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Aeomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 14440
; LENGTH: 392
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC003098.1
; OTHER INFORMATION: EXPRESSED IN BONE, SIGNAL = 0.62
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.49
US-09-864-761-14440

Query Match 10.8%; Score 69; DB 10; Length 392;
Best Local Similarity 100.0%; Pred.No. 7.5e-27;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy 475 CTGTGGCCCTCGTGCAGTGCAGCGCCTCACCCGGTTCACCAACAGTCGGAGCTCAAG 534
Db 280 CTGTGGCCCTCGTGCAGTGCAGCGCCTCACCCGGTTCACCAACAGTCGGAGCTCAAG 221

Qy 535 GACTTCGGG 543
Db 220 GACTTCGGG 212

RESULT 3
US-09-864-761-30988/c
; Sequence 30988, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30

```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2003, 01:33:30 ; Search time 76.5935 Seconds
(without alignments)
7066.115 Million cell updates/sec

Title: US-09-867-274-3

Perfect score: 636

Sequence: 1 atgcgacctactagcccc.....agctggagaacgcctactag 636

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 574371 seqs, 425486471 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1148742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	636	100.0	636	10	US-09-867-274-3	Sequence 3, Appli
C 2	69	10.8	392	10	US-09-864-761-14440	Sequence 14440, A
C 3	69	10.8	422	10	US-09-864-761-30988	Sequence 30988, A
C 4	69	10.8	759	10	US-09-867-274-1	Sequence 1, Appli
C 5	26	4.1	41	10	US-09-867-274-20	Sequence 20, Appl
C 6	25	3.9	25	10	US-09-867-274-14	Sequence 14, Appl
C 7	25	3.9	45	10	US-09-867-274-19	Sequence 19, Appl
C 8	24	3.8	24	10	US-09-867-274-12	Sequence 12, Appl
C 9	21	3.3	21	10	US-09-867-274-15	Sequence 15, Appl
C 10	21	3.3	21	10	US-09-867-274-17	Sequence 17, Appl
C 11	21	3.3	29	10	US-09-867-274-9	Sequence 9, Appli
C 12	21	3.3	577	10	US-09-864-761-9695	Sequence 9695, Ap
C 13	20	3.1	24	10	US-09-867-274-10	Sequence 10, Appl
C 14	19	3.0	198285	10	US-09-880-107-3814	Sequence 3814, Ap
C 15	18	2.8	323	10	US-09-867-550-675	Sequence 675, App
C 16	18	2.8	911	7	US-08-726-211-6	Sequence 6, Appli
C 17	18	2.8	5086	7	US-08-726-211-4	Sequence 4, Appli
C 18	18	2.8	6011	9	US-09-887-527-34	Sequence 34, Appl
C 19	18	2.8	24757	9	US-09-764-868-1347	Sequence 1347, Ap

ALIGNMENTS

RESULT 1
US-09-867-274-3
; Sequence 3, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 3
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-867-274-3

Query Match 100.0%; Score 636; DB 10; Length 636;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCAGCCCTCCTAGCCCGTGCCTCATCTGCTACTTGTGCACGCTGCTTCTGTGCT 60
Db 1 ATGCAGCCCTCCTAGCCCGTGCCTCATCTGCTACTTGTGCACGCTGCTTCTGTGCT 60
QY 61 GTGAGGGCCAGGGGTGGCAAGCTTTCAGGAATGATGCCACAGAGGTTCATCCAGGGCTT 120
Db 61 GTGAGGGCCAGGGGTGGCAAGCTTTCAGGAATGATGCCACAGAGGTTCATCCAGGGCTT 120
QY 121 GGAGAGTACCCCGAGCTCTCTCTGAGAACCAACAGACCATGAAACCGGGCGGAGATGGA 180
Db 121 GGAGAGTACCCCGAGCTCTCTCTGAGAACCAACAGACCATGAAACCGGGCGGAGATGGA 180
QY 181 GGCAGACCTCCCAACCATCCCTATGACCCCAAGATGTGTCGAGTACAGCTCCCGGAG 240
Db 181 GGCAGACCTCCCAACCATCCCTATGACCCCAAGATGTGTCGAGTACAGCTCCCGGAG 240

; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 20
; LENGTH: 19
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer for PCR
US-09-449-218D-20

Query Match 3.0%; Score 19; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 303 GGTGTGCTCGGCCAGTGC 321
|||
Db 19 GGTGTGCTCGGCCAGTGC 1

RESULT 13
US-09-449-218D-25/c
; Sequence 25, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 25
; LENGTH: 19
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer for PCR
US-09-449-218D-25

Query Match 3.0%; Score 19; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 303 GGTGTGCTCGGCCAGTGC 321
|||
Db 19 GGTGTGCTCGGCCAGTGC 1

RESULT 14
US-09-449-218D-29/c
; Sequence 29, Application US/09449218D
; Patent No. 6395511

; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 29
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer for PCR
US-09-449-218D-29

Query Match 2.8%; Score 18; DB 4; Length 29;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 619 CTGGAGAACGCCTACTAG 636
|||||
Db 29 CTGGAGAACGCCTACTAG 12

RESULT 15
US-09-449-218D-38
; Sequence 38, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 38
; LENGTH: 31
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer for PCR
US-09-449-218D-38

Query Match 2.8%; Score 18; DB 4; Length 31;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 248 ACACCCGCTTCTGACAG 265
|||||
Db 14 ACACCCGCTTCTGACAG 31

Search completed: March 29, 2003, 10:47:09
Job time : 55.2086 secs

; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; TITLE OF INVENTION: BONE MINERALIZATION
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 2301
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-449-218D-7

Query Match 10.8%; Score 69; DB 4; Length 2301;
Best Local Similarity 100.0%; Pred. No. 7e-24;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 475 CTGGTGGCTCTGCAAGTGAAGCGCTTCCACCAACAGTCGGAGCTCAAG 534
Db 528 CTGGTGGCTCTGCAAGTGAAGCGCTTCCACCAACAGTCGGAGCTCAAG 587

QY 535 GACTTCGGG 543
Db 588 GACTTCGGG 596

RESULT 9

US-09-449-218D-18
; Sequence 18, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; TITLE OF INVENTION: BONE MINERALIZATION
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 9301
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-449-218D-18

Query Match 10.8%; Score 69; DB 4; Length 9301;
Best Local Similarity 100.0%; Pred. No. 6.4e-24;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 475 CTGGTGGCTCTGCAAGTGAAGCGCTTCCACCAACAGTCGGAGCTCAAG 534
Db 3447 CTGGTGGCTCTGCAAGTGAAGCGCTTCCACCAACAGTCGGAGCTCAAG 3506

QY 535 GACTTCGGG 543
Db 3507 GACTTCGGG 3515

RESULT 10

US-09-449-218D-9

; Sequence 9, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; TITLE OF INVENTION: BONE MINERALIZATION
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 642
; TYPE: DNA
; ORGANISM: Cercopithecus pygerythrus
US-09-449-218D-9

Query Match 10.7%; Score 68; DB 4; Length 642;
Best Local Similarity 100.0%; Pred. No. 2.3e-23;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 475 CTGGTGGCTCTGCAAGTGAAGCGCTTCCACCAACAGTCGGAGCTCAAG 534
Db 481 CTGGTGGCTCTGCAAGTGAAGCGCTTCCACCAACAGTCGGAGCTCAAG 540

QY 535 GACTTCGGG 542
Db 541 GACTTCGGG 548

RESULT 11

US-09-449-218D-15
; Sequence 15, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; TITLE OF INVENTION: BONE MINERALIZATION
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 532
; TYPE: DNA
; ORGANISM: Bos torus
US-09-449-218D-15

Query Match 5.0%; Score 32; DB 4; Length 532;
Best Local Similarity 100.0%; Pred. No. 3.2e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 475 CTGGTGGCTCTGCAAGTGAAGCGCTTCCACCAACAGTCGGAGCTCAAG 506
Db 390 CTGGTGGCTCTGCAAGTGAAGCGCTTCCACCAACAGTCGGAGCTCAAG 421

RESULT 12

US-09-449-218D-20/c
; Sequence 20, Application US/09449218D

```
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 45
; LENGTH: 642
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(639)
US-09-449-218D-45

Query Match          10.8%; Score 69; DB 4; Length 642;
Best Local Similarity 100.0%; Pred. No. 7e-24;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 475 CTGGTGGCCTCGTGCAGAGTCAAGCGCCTCACCGCTTCCACAAACAGTCGGAGCTCAAG 534
    |||||||
Db 481 CTGGTGGCCTCGTGCAGAGTCAAGCGCCTCACCGCTTCCACAAACAGTCGGAGCTCAAG 540
    |||||||

QY 535 GACTTCGGG 543
    |||||||
Db 541 GACTTCGGG 549

RESULT 5
US-09-449-218D-1
; Sequence 1, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2301
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-449-218D-1

Query Match          10.8%; Score 69; DB 4; Length 2301;
Best Local Similarity 100.0%; Pred. No. 7e-24;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 475 CTGGTGGCCTCGTGCAGAGTCAAGCGCCTCACCGCTTCCACAAACAGTCGGAGCTCAAG 534
    |||||||
Db 528 CTGGTGGCCTCGTGCAGAGTCAAGCGCCTCACCGCTTCCACAAACAGTCGGAGCTCAAG 587
    |||||||

QY 535 GACTTCGGG 543
    |||||||
Db 588 GACTTCGGG 596

RESULT 6
US-09-449-218D-3
; Sequence 3, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
```

```
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; TITLE OF INVENTION: BONE MINERALIZATION
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 2301
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-449-218D-3

Query Match          10.8%; Score 69; DB 4; Length 2301;
Best Local Similarity 100.0%; Pred. No. 7e-24;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 475 CTGGTGGCCTCGTGCAGAGTCAAGCGCCTCACCGCTTCCACAAACAGTCGGAGCTCAAG 534
    |||||||
Db 528 CTGGTGGCCTCGTGCAGAGTCAAGCGCCTCACCGCTTCCACAAACAGTCGGAGCTCAAG 587
    |||||||

QY 535 GACTTCGGG 543
    |||||||
Db 588 GACTTCGGG 596

RESULT 7
US-09-449-218D-5
; Sequence 5, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; TITLE OF INVENTION: BONE MINERALIZATION
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 2301
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-449-218D-5

Query Match          10.8%; Score 69; DB 4; Length 2301;
Best Local Similarity 100.0%; Pred. No. 7e-24;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 475 CTGGTGGCCTCGTGCAGAGTCAAGCGCCTCACCGCTTCCACAAACAGTCGGAGCTCAAG 534
    |||||||
Db 528 CTGGTGGCCTCGTGCAGAGTCAAGCGCCTCACCGCTTCCACAAACAGTCGGAGCTCAAG 587
    |||||||

QY 535 GACTTCGGG 543
    |||||||
Db 588 GACTTCGGG 596

RESULT 8
US-09-449-218D-7
; Sequence 7, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
```


Db 301 TTGGTGTCTCGGCAAGTGGGCCCCCGGGGCTGCTGCCAAGCCATCGGGCGGTG 360
Qy 361 AAGTGGTGGCGCCGAAACGAGCCGATTTCCGCTGATCCCGGATCGCTACCGCGCGAG 420
Db 361 AAGTGGTGGCGCCGAAACGAGCCGATTTCCGCTGATCCCGGATCGCTACCGCGCGAG 420
Qy 421 CGGGTGCAGTCTGTGCCCCCGGGGCGGGCGCGCTGCGCAAGTGTGCTGTGGT 480
Db 421 CGGGTGCAGTCTGTGCCCCCGGGGCGGGCGCGCTGCGCAAGTGTGCTGTGGT 480
Qy 481 GCTCTGTGCAAGTCAAGCGCTTCCACCGCTTCCACACAGTTCGAGCTCAAGGACTTC 540
Db 481 GCTCTGTGCAAGTCAAGCGCTTCCACCGCTTCCACACAGTTCGAGCTCAAGGACTTC 540
Qy 541 GGGCGGAGACCGCGGCGCGGAGAGGTCGCAAGCGCGCGCGCGCGCGCGCGCGCGCG 600
Db 541 GGGCGGAGACCGCGGCGCGGAGAGGTCGCAAGCGCGCGCGCGCGCGCGCGCGCGCG 600
Qy 601 AAGCCCAACGAGCGAGCTGGAGACGCTACTAG 636
Db 601 AAGCCCAACGAGCGAGCTGGAGACGCTACTAG 636

RESULT 2

US-09-449-218D-17
; Sequence 17, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 35828
; TYPE: DNA
; ORGANISM: Mus musculus
; NAME/KEY: misc feature
; LOCATION: (1)-(35828)
; OTHER INFORMATION: n = A,T,C or G
US-09-449-218D-17

Query Match 66.7%; Score 424; DB 4; Length 35828;
Best Local Similarity 100.0%; Pred. No. 5.1e-193;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 213 AGATGTGTCGAGTACAGTGGCGGAGTGCATACACCGCTTCTGACAGCGGCC 272
Db 21021 AGATGTGTCGAGTACAGTGGCGGAGTGCATACACCGCTTCTGACAGCGGCC 21080
Qy 273 ATGCCCGAGCGCCAAAGCCGATCCAGGTTGGTGTGCTCCGGCCAGTGGCGCCCGCGCG 332
Db 21081 ATGCCCGAGCGCCAAAGCCGATCCAGGTTGGTGTGCTCCGGCCAGTGGCGCCCGCGCG 21140
Qy 333 GTGTGTCGCAACCGCATCGGGCGCGTGAAGTGGTGGCGCCGAAACGAGATTTCCG 392
Db 21141 GTGTGTCGCAACCGCATCGGGCGCGTGAAGTGGTGGCGCCGAAACGAGATTTCCG 21200
Qy 393 CTGCATCCCGGATCGTACCGCGCGAGGGTGCAGTGTGCTGCGCGGGCGCGCG 452
Db 21201 CTGCATCCCGGATCGTACCGCGCGAGGGTGCAGTGTGCTGCGCGGGCGCGCG 21260

Qy 453 GCGCGCTCGCGCAAGTGGCTCTGGTGGCTCGTCAAGTGAAGCGCTCACCGCTT 512
Db 21261 GCGCGCTCGCGCAAGTGGCTCTGGTGGCTCGTCAAGTGAAGCGCTCACCGCTT 21320
Qy 513 CCACCAACAGTCCGAGCTCAAGGACTTCGGGCGGAGACCGCGCGCGCGAGAGGGTGC 572
Db 21321 CCACCAACAGTCCGAGCTCAAGGACTTCGGGCGGAGACCGCGCGCGCGAGAGGGTGC 21380
Qy 573 CAAGCTA 632
Db 21381 CAAGCTA 21440
Qy 633 CTAG 636
Db 21441 CTAG 21444

RESULT 3

US-09-449-218D-13
; Sequence 13, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 674
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-09-449-218D-13

Query Match 16.4%; Score 104; DB 4; Length 674;
Best Local Similarity 99.4%; Pred. No. 1.6e-40;
Matches 154; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 388 TTCCGCTGCATCCCGGATCCGATCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 447
Db 426 TTCCGCTGCATCCCGGATCCGATCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 485
Qy 448 GCGGCGCGCGCGCTCGCGCAAGGTGCGTCTGGTGGCTCGTGAAGTGAAGCGCTCAC 507
Db 486 GCGGCGCGCGCGCTCGCGCAAGGTGCGTCTGGTGGCTCGTGAAGTGAAGCGCTCAC 545
Qy 508 CGTTCCACCAACAGTCCGAGCTCAAGGACTTCGG 542
Db 546 CGTTCCACCAACAGTCCGAGCTCAAGGACTTCGG 580

RESULT 4

US-09-449-218D-45
; Sequence 45, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 45
; LENGTH: 674
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-09-449-218D-45

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2003, 01:31:05 ; Search time 39.2086 Seconds
(without alignments)
4974.582 Million cell updates/sec

Title: US-09-867-274-3

Perfect score: 636

Sequence: 1 atgcagccctactagcccc.....agctggagaagcctactag 636

Scoring table:

OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 153338381 residues

Word size : 0

Total number of hits satisfying chosen parameters: 892724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/1/ina/PTUS_COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	585	92.0	638	4	US-09-449-218D-11
2	424	66.7	35828	4	US-09-449-218D-17
3	104	16.4	874	4	US-09-449-218D-13
4	69	10.8	842	4	US-09-449-218D-45
5	69	10.8	2301	4	US-09-449-218D-1
6	69	10.8	2301	4	US-09-449-218D-3
7	69	10.8	2301	4	US-09-449-218D-5
8	69	10.8	2301	4	US-09-449-218D-7
9	69	10.8	2301	4	US-09-449-218D-9
10	68	10.7	842	4	US-09-449-218D-15
11	32	5.0	532	4	US-09-449-218D-15
12	19	3.0	19	4	US-09-449-218D-20
13	19	3.0	19	4	US-09-449-218D-25
14	18	2.8	29	4	US-09-449-218D-29
15	18	2.8	31	4	US-09-449-218D-38
16	18	2.8	57	4	US-09-449-218D-27
17	18	2.8	334	4	US-09-060-756-261
18	18	2.8	554	2	US-08-738-127-2
19	18	2.8	615	2	US-08-465-485A-22
20	18	2.8	615	3	US-09-080-285-22
21	18	2.8	623	6	5506344-3
22	18	2.8	717	2	US-08-465-485A-20
23	18	2.8	717	3	US-09-080-285-20
24	18	2.8	760	1	US-08-405-702A-11
25	18	2.8	831	6	5459251-3
26	18	2.8	831	6	5506344-4
27	18	2.8	911	4	US-09-126-109-5
					Sequence 11, Appl
					Sequence 17, Appl
					Sequence 45, Appl
					Sequence 1, Appl
					Sequence 3, Appl
					Sequence 5, Appl
					Sequence 7, Appl
					Sequence 18, Appl
					Sequence 9, Appl
					Sequence 15, Appl
					Sequence 20, Appl
					Sequence 25, Appl
					Sequence 29, Appl
					Sequence 38, Appl
					Sequence 27, Appl
					Sequence 2, Appl
					Sequence 22, Appl
					Sequence 22, Appl
					Patent No. 5506344
					Sequence 20, Appl
					Sequence 20, Appl
					Patent No. 5459251
					Patent No. 5506344
					Sequence 5, Appl

c 28 18 2.8 911 5 PCT-US93-06251-3 Sequence 3, Appli
29 18 2.8 1405 1 US-08-035-558-3 Sequence 3, Appli
30 18 2.8 1405 2 US-08-682-847-5 Sequence 5, Appli
31 18 2.8 1405 4 US-09-063-676-1 Sequence 1, Appli
32 18 2.8 1416 3 US-08-911-853-3 Sequence 3, Appli
33 18 2.8 1416 4 US-09-479-409-3 Sequence 3, Appli
34 18 2.8 1416 4 US-09-479-453-3 Sequence 3, Appli
35 18 2.8 1846 2 US-08-365-486A-16 Sequence 16, Appl
36 18 2.8 1846 4 US-08-880-342-16 Sequence 16, Appl
37 18 2.8 4377 3 US-08-911-853-28 Sequence 28, Appl
38 18 2.8 4377 4 US-09-479-409-28 Sequence 28, Appl
39 18 2.8 4377 4 US-09-479-453-28 Sequence 28, Appl
40 18 2.8 4749 1 US-08-452-259-1 Sequence 1, Appli
41 18 2.8 4749 5 PCT-US96-07336-1 Patent No. 5459251
42 18 2.8 4825 6 5459251-1 Sequence 19, Appl
43 18 2.8 5086 2 US-08-465-485A-19 Sequence 14, Appl
44 18 2.8 5086 2 US-08-365-486A-14 Sequence 19, Appl
45 18 2.8 5086 3 US-09-080-285-19

ALIGNMENTS

RESULT 1
US-09-449-218D-11
; Sequence 11, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; TITLE OF INVENTION: BONE MINERALIZATION
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 638
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-449-218D-11

Query Match 92.0%; Score 585; DB 4; Length 638;
Best Local Similarity 99.8%; Pred. No. 1.4e-269;
Matches 635; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGCAGCCCTCAGTCCCGCTGCTCTCATCTGCTACTGCTGACGCTGCTTCTGCT 60
Db 1 ATGCAGCCCTCAGTCCCGCTGCTCTCATCTGCTACTGCTGACGCTGCTTCTGCT 60
Qy 61 GTGAGGCGCCAGGGGTGGCAAGCTTCAGGAATGATGCCACAGAGGTCTATCCAGGGCTT 120
Db 61 GTGAGGCGCCAGGGGTGGCAAGCTTCAGGAATGATGCCACAGAGGTCTATCCAGGGCTT 120
Qy 121 GGAGAGTACCCCGAGCTCTCTCTGAGAA CAACAGACCATGAACCCGGCGGAGATGGA 180
Db 121 GGAGAGTACCCCGAGCTCTCTCTGAGAA CAACAGACCATGAACCCGGCGGAGATGGA 180
Qy 181 GGACAGCTCCCGACCATCTCTATGAGCCCAAGATGTGTCGAGTACAGTCCCGCGAG 240
Db 181 GGACAGCTCCCGACCATCTCTATGAGCCCAAGATGTGTCGAGTACAGTCCCGCGAG 240
Qy 241 CTGCACCTACACCCCGCTTCTTGACAGACGGCCCATGCGGAGCGCCAGCGGTACCGAG 300
Db 241 CTGCACCTACACCCCGCTTCTTGACAGACGGCCCATGCGGAGCGCCAGCGGTACCGAG 300
Qy 301 TTGTGTGCTCCGCCAGTGGCGGCCCGCGGCTGTCTGCCCAACGCCCATCGGCGCGTG 360

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 105 GGTCATCCCGGGCTTGGAGAG 126
| | | | | | | | | | | | | | | | | |
Db 40 GGTCATCCCGGGCTTGGAGAG 61

Search completed: March 29, 2003, 10:45:09
Job time : 1374.57 secs

RPCI-24 Mouse BAC Library produced by Pieter de Jong. The library was cloned in the pTARBAC1 cloning vector at the BamHI sites using MboI partially digested male C57BL/6J DNA."

BASE COUNT 30 a 44 c 45 g 28 t
ORIGIN

Query Match 7.1%; Score 45; DB 17; Length 147;
Best Local Similarity 100.0%; Pred. No. 4e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 222 CGAGTACACTCCGGAGCTGACCTACACCGCTTCTTGACAGA 266
|||||
Db 101 CGAGTACACTCCGGAGCTGACCTACACCGCTTCTTGACAGA 145
|||||

RESULT 14
BF523030 254 bp mRNA linear EST 11-DEC-2000
LOCUS
DEFINITION UI-R-C2p-rk-a-03-0-UI-r1 UI-R-C2p Rattus norvegicus cDNA clone
UI-R-C2p-rk-a-03-0-UI 5', mRNA sequence.

ACCESSION BF523030
VERSION BF523030.1 GI:11631045
KEYWORDS EST.

SOURCE Norway rat.

ORGANISM

Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

1 (bases 1 to 254)

REFERENCE

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene

discovery

JOURNAL Genome Res.

MEDLINE 6 (9), 791-806 (1996)

COMMENT 97044477

Contact: Soares, MB

Program for Rat Gene Discovery and Mapping

University of Iowa

451 Eckstein Medical Research Building Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: msoares@blue.weeg.uiowa.edu

cDNA Library Preparation: M.B. Soares Lab Clone distribution:

clones will be available through Research Genetics (www.resgen.com)

This clone is also available through the I.M.A.G.E. Consortium at

LLNL (info@image.llnl.gov). IMAGE ID= 1793134

Seq primer: M13 Forward.

Location/Qualifiers

1..254

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/clone="UI-R-C2p-rk-a-03-0-UI"

/clone_lib="UI-R-C2p"

/dev_stage="adult"

/lab_host="DH10B (Life Technologies)"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-C2p

library is a subtracted library derived from the UI-R-C1

library, which is a subtracted library derived from the

UI-R-C0 library. The UI-R-C0 library consisted of a

mixture of individually tagged normalized libraries

constructed from rat placenta, adult lung, brain, liver,

kidney, heart, spleen, ovary, muscle, 8, 12 and 18-day

embryo. The tag is a string of 3-5 nucleotides present

between the Not I site and the oligo-dr track which allows

identification of the library of origin of a clone within

the mixture. The subtracted library (UI-R-C2p) was

constructed as follows: PCR amplified cDNA inserts from

UI-R-C1 clones from which 3' ESTs had been derived was

used as a driver in a hybridization with the UI-R-C1

library in the form of single-stranded circles. The

remaining single-stranded circles (subtracted library) was

purified by hydroxyapatite column chromatography,
converted to double-stranded circles and electroporated
into DH10B bacteria (Life Technologies) to generate the
UI-R-C2p library. This procedure has been previously
described (Bonaldo, Lennon and Soares, Genome Research 6:
791-806, 1996).

BASE COUNT 61 a 50 c 53 g 50 t
ORIGIN

Query Match 3.6%; Score 23; DB 12; Length 254;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 146 AGAACACCAACGACCATGACCGG 168
|||||

Db 186 AGAACACCAACGACCATGACCGG 208
|||||

RESULT 15

BE835032

LOCUS

DEFINITION RC6-FN0019-290600-011-A02 FN0019 Homo sapiens cDNA, mRNA sequence.

ACCESSION BE835032

VERSION BE835032.1 GI:10267410

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 385)

REFERENCE

AUTHORS

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,

Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,

Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare

,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A.

MEDLINE 97 (7), 3491-3496 (2000)

COMMENT 20202663

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC6-FN0019-290

600-011-A02&t3=2000-06-29&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 14

High quality sequence stop: 361.

Location/Qualifiers

1..385

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_lib="FN0019"

/dev_stage="Adult"

/note="Organ: prostate normal; Vector: puc18; Site_1: SmaI

; Site_2: SmaI; A mini-library was made by cloning

products derived from ORESTES PCR (U.S. Letters Patent

application No. 196,716 - Ludwig Institute for Cancer

Research) profiles into the pUC 18 vector. Reverse

transcription of tissue mRNA and cDNA amplification were

performed under low stringency conditions."

92 a 89 c 119 g 85 t

BASE COUNT

ORIGIN

Query Match 3.5%; Score 22; DB 12; Length 385;

Best Local Similarity 100.0%; Pred. No. 6.9;

Soares Lab Clone distribution: clones will be available through Research Genetics (www.resgen.com) The following repetitive elements were found in this cdna sequence: 22-42,
>GC rich#Low complexity
Seq primer: M13 Forward
POLYA=No.

FEATURES

source

Location/Qualifiers

1. .472

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/clone="UI-R-BJ2-bgm-b-02-0-UI"

/clone_lib="UI-R-BJ2"

/lab_host="DH10B (Life Technologies)"

/notes="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-BJ2 library is a subtracted library derived from the following tissues: heart, atrium at 15 dpc, ventricle at 16.5 dpc, atrium at 16.5 dpc, ventricle at 13 dpc, ventricle at 15 dpc, AV canal at 15 dpc. For a detailed description of the library from which this clone was derived, please visit our web site at ratest.eng.uiowa.edu. The subtraction has been previously described in (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)
TAG_SEQ=None found"

103 a 148 c 147 g 74 t

BASE COUNT
ORIGIN

Query Match 7.9%; Score 50; DB 12; Length 472;

Best Local Similarity 100.0%; Pred. No. 1.7e-13;

Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 587 GCGCCGGGAGCCAAAGCCAGCGGAGCTGGAGAACGCTACTAG 636

Db 33 GCGCCGGGAGCCAAAGCCAGCGGAGCTGGAGAACGCTACTAG 82

RESULT 12

BF407514

LOCUS

DEFINITION UI-R-BJ2-bqt-b-08-0-UI.s1 UI-R-BJ2 Rattus norvegicus cDNA clone

UI-R-BJ2-bqt-b-08-0-UI 3', mRNA sequence.

ACCESSION BF407514

VERSION BF407514.1 GI:11395489

KEYWORDS EST.

SOURCE Norway rat.

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

REFERENCE 1 (bases 1 to 532)

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene

discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 97044477

COMMENT Contact: Soares, MB

Program for Rat Gene Discovery and Mapping

University of Iowa

451 Eckstein Medical Research Building Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: msoares@blue.weeg.uiowa.edu

Oligo-dT track not found, Not I site shown in beginning of sequence

is likely internal to the message. cDNA Library Preparation: M.B.

Soares Lab Clone distribution: clones will be available through

Research Genetics (www.resgen.com) The following repetitive

elements were found in this cdna sequence: 22-42,

>GC rich#Low complexity

Seq primer: M13 Forward

POLYA=No.

FEATURES

source

Location/Qualifiers

1. .532

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/clone="UI-R-BJ2-bqt-b-08-0-UI"

/clone_lib="UI-R-BJ2"

/lab_host="DH10B (Life Technologies)"

/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-BJ2 library is a subtracted library derived from the following tissues: heart, atrium at 15 dpc, ventricle at 16.5 dpc, atrium at 16.5 dpc, ventricle at 13 dpc, ventricle at 15 dpc, AV canal at 15 dpc. For a detailed description of the library from which this clone was derived, please visit our web site at ratest.eng.uiowa.edu. The subtraction has been previously described in (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)
TAG_SEQ=None found"

111 a 160 c 167 g 94 t

BASE COUNT
ORIGIN

Query Match 7.9%; Score 50; DB 12; Length 532;

Best Local Similarity 100.0%; Pred. No. 1.7e-13;

Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 587 GCGCCGGGAGCCAAAGCCAGCGGAGCTGGAGAACGCTACTAG 636

Db 33 GCGCCGGGAGCCAAAGCCAGCGGAGCTGGAGAACGCTACTAG 82

RESULT 13

AZ746508

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AZ746508 147 bp DNA linear GSS 25-JAN-2001
RPCI-24-79P16.TJB RPCI-24 Mus musculus genomic clone RPCI-24-79P16,
DNA sequence.

AZ746508

AZ746508.1 GI:12529424

GSS.

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 147)

Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M.,

Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E.,

Russell,D., de Jong,P. and Fraser,C.M.

Mouse BAC End Sequences from Library RPCI-24

Unpublished (1999)

Other_GSSs: RPCI-24-79P16.TVB

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-24. For BAC

library availability, please contact Pieter de Jong

(pdejong@mail.cho.org). Clones may be purchased from BACPACResources (<http://www.chori.org/bacpac/orderingframe.htm>). BAC endpage: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html

Plate: 79 row: P column: 16

Seq primer: SP6

Class: BAC ends.

Location/Qualifiers

1. .147

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="RPCI-24-79P16"

/clone_lib="RPCI-24"

/sex="Male"

/cell_type="Spleen/Brain"

/notes="Vector: pTARBAC1; Site_1: BamHI; Site_2: BamHI;

Plate: 0191 row: D column: 21
Seq primer: CGTTGTAAACACGCGCCAGT
Class: plasmid ends
High quality sequence stop: 627.
Location/Qualifiers

FEATURES

source

```

1. .627
/organism="Mus musculus"
/strain="C57BL/6J"
/db xref="taxon:10090"
/clone="UUC2M0191D21"
/clone_lib="Mouse 10kb plasmid UUC1M library"
/sex="Male"
/lab host="E. Coli strain XL10-Gold, T1-resistant, F-"
/notes="Vector: pMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT      168 a 157 c 176 g 126 t
ORIGIN

```

```

Query Match      10.7%; Score 68; DB 17; Length 627;
Best Local Similarity 100.0%; Pred. No. 3.1e-22;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 213 AGATGTGTCGAGTACAGTGCAGCGAGCTGCACCTACACCGCTTCTGACAGCGGCC 272
      |||||||
Db 560 AGATGTGTCGAGTACAGTGCAGCGAGCTGCACCTACACCGCTTCTGACAGCGGCC 619

QY 273 ATGCCGCA 280
      |||||||
Db 620 ATGCCGCA 627

```

```

RESULT 10
AW434705
LOCUS
DEFINITION
  UI-R-BJ0p-afz-c-02-0-UI.s1 UI-R-BJ0p Rattus norvegicus cDNA clone
  UI-R-BJ0p-afz-c-02-0-UI 3', mRNA sequence.
ACCESSION
  AW434705
VERSION
  AW434705.1 GI:6966012
KEYWORDS
  EST.
SOURCE
  Norway rat.
  Rattus norvegicus

```

```

  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
  Rattus.
REFERENCE
  1 (bases 1 to 385)
  Bonaldo,M.F., Lennon,G. and Soares,M.B.
  Normalization and subtraction: two approaches to facilitate gene
  discovery
  Genome Res. 6 (9), 791-806 (1996)
  9704477
JOURNAL
  MEDLINE
  COMMENT
    Contact: Soares, MB
    Program for Rat Gene Discovery and Mapping
    University of Iowa
    451 Eckstein Medical Research Building Iowa City, IA 52242, USA
    Tel: 319 335 8250
    Fax: 319 335 9565

```

Email: msoares@blue.weeg.uiowa.edu
Oligo-dT track not found, Not 1 site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.B.
Soares Lab Clone distribution: clones will be available through
Research Genet:cs (www.resgen.com) The following repetitive
elements were found in this cDNA sequence: 21-41,
>GC-rich#Low complexity
Seq primer: M13 Forward
POLYA=No.

FEATURES

source

```

1. .305
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db xref="taxon:10116"
/clone="UI-R-BJ0p-afz-c-02-0-UI"
/clone_lib="UI-R-BJ0p"
/dev stage="adult"
/lab host="DH10B (Life Technologies)"
/notes="Vector: p773D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not 1; Site 2: Eco RI; The UI-R-BJ0p
library is a subtracted library derived from the UI-R-AA1,
UI-R-AB1, UI-R-AC1, UI-R-AD1, UI-R-AE1, UI-R-AF1, and
UI-R-AG1 libraries. These libraries represent tissues from
rat atrium at 16.5 dpc, ventricle at 16.5 dpc, AV canal
at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc, AV
canal at 15 dpc, and ventricle at 13 dpc. The tag is a
string of 5-6 nucleotides present between the Not I site
and the oligo-dT track. The library was constructed as
described by Bonaldo, Lennon and Soares, Genome Research
6: 791-806, 1996.
TAG L.B=UI-R-BJ0p
TAG T.SSUE=AV canal at 16.5 dpc
TAG SEQ=GAACC"
BASE COUNT      75 a 137 c 122 g 61 t
ORIGIN

```

```

Query Match      7.9%; Score 50; DB 10; Length 385;
Best Local Similarity 100.0%; Pred. No. 1.7e-13;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 587 GCGCCCGGAGCCAAAGCCAGCGAGCTGGAGAACGCTACTAG 636
      |||||||
Db 32 GCGCCCGGAGCCAAAGCCAGCGAGCTGGAGAACGCTACTAG 81

```

```

RESULT 11
BF420024
LOCUS
DEFINITION
  UI-R-BJ2-bpm-b-02-0-UI.s1 UI-R-BJ2 Rattus norvegicus cDNA clone
  UI-R-BJ2-bpm-b-02-0-UI 3', mRNA sequence.
ACCESSION
  BF420024
VERSION
  BF420024.1 GI:11408013
KEYWORDS
  EST.
SOURCE
  Norway rat.
  Rattus norvegicus

```

```

  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
  Rattus.
REFERENCE
  1 (bases 1 to 472)
  Bonaldo,M.F., Lennon,G. and Soares,M.B.
  Normalization and subtraction: two approaches to facilitate gene
  discovery
  Genome Res. 6 (9), 791-806 (1996)
  9704477
JOURNAL
  MEDLINE
  COMMENT
    Contact: Soares, MB
    Program for Rat Gene Discovery and Mapping
    University of Iowa
    451 Eckstein Medical Research Building Iowa City, IA 52242, USA
    Tel: 319 335 8250
    Fax: 319 335 9565

```

Email: msoares@blue.weeg.uiowa.edu
Oligo-dT track not found, Not 1 site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.B.

library is a subtracted library derived from the UI-R-C1 library, which is a subtracted library derived from the UI-R-CO library. The UI-R-CO library consisted of a mixture of individually tagged normalized libraries constructed from rat placenta, adult lung, brain, liver, kidney, heart, spleen, ovary, muscle, 8, 12 and 18-day embryo. The tag is a string of 3-5 nucleotides present between the Not I site and the oligo-dT track which allows identification of the library of origin of a clone within the mixture. The subtracted library (UI-R-C2p) was constructed as follows: PCR amplified cDNA inserts from UI-R-C1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with the UI-R-C1 library in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10 bacteria (Life Technologies) to generate the UI-R-C2p library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996)"

canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc, AV canal at 15 dpc, ventricle at 13 dpc, and adult heart. For a detailed description of the library from which this clone was derived, please visit our web site at ratseq.eng.uiowa.edu. The subtraction has been previously described in (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)
TAG_SEQ=None found"

BASE COUNT 41 a 101 c 109 g 40 t
ORIGIN

Query Match 16.4%; Score 104; DB 10; Length 291;
Best Local Similarity 99.4%; Pred. No. 8.7e-40;
Matches 154; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 388 TTCGGCTGATCCCGGATCGCTACCGCGCGAGCGGTGCGAGCTGCTGTGCCCGCGGGGC 447
|||
DB 174 TTCGGCTGATCCCGGATCGCTACCGCGCGAGCGGTGCGAGCTGCTGTGCCCGCGGGGC 115
|||
QY 448 GCGGCGCGCGCTCGCGCAAGTGGCTGTGGTGGCTGTGCAAGTGAAGCGCTCACC 507
|||
DB 114 GCGGCGCGCGCTCGCGCAAGTGGCTGTGGTGGCTGTGCAAGTGAAGCGCTCACC 55
|||
QY 508 CGCTTCCCAACAGTCGGAGCTCAAGGACTTCGG 542
|||
DB 54 CGCTTCCCAACAGTCGGAGCTCAAGGACTTCGG 20

RESULT 6

BE101082/c
LOCUS BE101082 419 bp mRNA linear EST 13-JUN-2000
DEFINITION UI-R-BJ1-aty-h-10-0-UI.s1 UI-R-BJ1 Rattus norvegicus cDNA clone
UI-R-BJ1-aty-h-10-0-UI 3', mRNA sequence.

ACCESSION BE101082
VERSION BE101082.1 GI:8492989

KEYWORDS EST.

SOURCE Norway rat.

ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 419)

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene

discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 9704477

COMMENT

Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565

Email: msoares@blue.weeg.uiowa.edu
Oligo-dT track not found. Not I site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.B.
Soares Lab Clone distribution: Clones will be available through
Research Genetics (www.resgen.com) The following repetitive
elements were found in this cDNA sequence: 97-125,
>GC rich#Low complexity

Seq primer: M13 Forward

POLYA=No.

FEATURES

source

1..419

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/clone="UI-R-BJ1-aty-h-10-0-UI"

/clone_lib="UI-R-BJ1"

/lab_host="DH10B (Life Technologies)"

/note="Vector: pRT3D-Pac (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-BJ1
library is a subtracted library derived from the following

tissues: atrium at 16.5 dpc, ventricle at 16.5 dpc, AV canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc, AV canal at 15 dpc, ventricle at 13 dpc, and adult heart. For a detailed description of the library from which this clone was derived, please visit our web site at ratseq.eng.uiowa.edu. The subtraction has been previously described in (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)
TAG_SEQ=None found"

BASE COUNT 56 a 135 c 155 g 73 t
ORIGIN

Query Match 16.4%; Score 104; DB 10; Length 419;
Best Local Similarity 99.4%; Pred. No. 9.1e-40;
Matches 154; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 388 TTCGGCTGATCCCGGATCGCTACCGCGCGAGCGGTGCGAGCTGCTGTGCCCGCGGGGC 447
|||
DB 174 TTCGGCTGATCCCGGATCGCTACCGCGCGAGCGGTGCGAGCTGCTGTGCCCGCGGGGC 115
|||
QY 448 GCGGCGCGCGCTCGCGCAAGTGGCTGTGGTGGCTGTGCAAGTGAAGCGCTCACC 507
|||
DB 114 GCGGCGCGCGCTCGCGCAAGTGGCTGTGGTGGCTGTGCAAGTGAAGCGCTCACC 55
|||
QY 508 CGCTTCCCAACAGTCGGAGCTCAAGGACTTCGG 542
|||
DB 54 CGCTTCCCAACAGTCGGAGCTCAAGGACTTCGG 20

RESULT 7

AI556282/c
LOCUS AI556282 360 bp mRNA linear EST 23-MAR-1999
DEFINITION UI-R-C2p-rk-a-03-0-UI.s1 UI-R-C2p Rattus norvegicus cDNA clone
UI-R-C2p-rk-a-03-0-UI 3', mRNA sequence.

ACCESSION AI556282

VERSION AI556282.1 GI:4488645

KEYWORDS EST.

SOURCE Norway rat.

ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 360)

AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE Normalization and subtraction: two approaches to facilitate gene

discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

MEDLINE 9704477

COMMENT

Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565

Email: msoares@blue.weeg.uiowa.edu
Oligo-dT track not found. Not I site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.
Patina Bonaldo, Ph.D. Clone distribution: Clones will be available
through Research Genetics (www.resgen.com) The following repetitive
elements were found in this cDNA sequence: 97-125,
>GC rich#Low complexity

Seq primer: M13 Forward

Locat:on/Qualifiers

1..360

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/clone="UI-R-C2p-rk-a-03-0-UI"

/clone_lib="UI-R-C2p"

/dev_stage="adult"

/lab_host="DH10B (Life Technologies)"

/note="Vector: pRT3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-C2p

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2003, 01:29:00 ; Search time 1369.57 Seconds
(without alignments)
7520.874 Million cell updates/sec

Title: US-09-867-274-3

Perfect score: 636

Sequence: 1 atgcagccctactagcccc.....agctggagaacgcctactag 636

Scoring table: OLIGO_NUC

Gapop_60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

```

EST: *
1: em_estba: *
2: em_esthum: *
3: em_estin: *
4: em_estmu: *
5: em_estov: *
6: em_estpl: *
7: em_estro: *
8: em_htc: *
9: gb_est1: *
10: gb_est2: *
11: gb_htc: *
12: gb_est3: *
13: gb_est4: *
14: gb_est5: *
15: em_estum: *
16: em_estom: *
17: gb_gss: *
18: em_gss_hum: *
19: em_gss_inv: *
20: em_gss_pln: *
21: em_gss_vrt: *
22: em_gss_fun: *
23: em_gss_nam: *
24: em_gss_mus: *
25: em_gss_other: *
26: em_gss_pro: *
27: em_gss_rtd: *

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	636	100.0	1990	11 AK017295	AK017295 Mus muscu
2	583	91.7	623	10 BB637315	BB637315 BB637315
3	469	73.7	511	10 BB638050	BB638050 BB638050
4	350	55.0	667	10 BB636457	BB636457 BB636457
5	104	16.4	291	10 BB112224	BB112224 UI-R-BJ1-
6	104	16.4	419	10 BB101082	BB101082 UI-R-BJ1-

C	7	98	15.4	360	9	AI556282
C	8	71	11.2	315	9	AI556282
C	9	68	10.7	627	17	AZ876400
	10	50	7.9	385	10	AZ876400
	11	50	7.9	472	12	AF420024
	12	50	7.9	532	12	BF407514
	13	45	7.1	147	17	AZ746508
	14	23	3.6	254	12	BF523030
	15	22	3.5	385	12	BE835032
	16	22	3.5	535	17	AQ171546
	17	21	3.3	301	10	BB229435
	18	21	3.3	315	9	AI556282
	19	21	3.3	351	10	BB115441
	20	21	3.3	422	13	BM668009
	21	3.3	458	9	AI096436	
	22	21	3.3	460	14	BM689096
	23	21	3.3	477	10	AQ233881
	24	21	3.3	497	17	AQ266759
	25	21	3.3	552	14	BM696978
	26	21	3.3	646	17	CNS05C5X
	27	21	3.3	787	12	BG820544
	28	20	3.1	248	12	BG050913
	29	20	3.1	249	17	AZ576934
	30	20	3.1	294	10	AW335927
	31	20	3.1	296	10	AW335370
	32	20	3.1	298	10	AW335931
	33	20	3.1	426	13	BI537845
	34	20	3.1	434	9	AA974943
	35	20	3.1	537	10	AW461827
	36	20	3.1	622	12	BG851121
	37	20	3.1	935	13	BM451086
	38	19	3.0	243	14	BQ469230
	39	19	3.0	259	12	BG604260
	40	19	3.0	282	10	BB386370
	41	19	3.0	308	10	BB205564
	42	19	3.0	344	10	BB626832
	43	19	3.0	328	10	BB536166
	44	19	3.0	424	12	BF605457
	45	19	3.0	432	14	R68738

ALIGNMENTS

RESULT 1
AK017295
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE

AK017295
Mus musculus 6 days neonate head cDNA, RIKEN full-length enriched library, clone:5430411E23:sclerostin, full insert sequence.
AK017295
HTC; CAP trapper.
Mus musculus (strain:C57BL/6J) 6 days neonate head cDNA to mRNA, clone lib:RIKEN full-length enriched mouse cDNA library
Clone:5430411E23.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3

RESULT 15

AAA29061
ID AAA29061 standard; cDNA; 2301 BP.

XX AC AAA29061;

XX DT 12-SEP-2000 (first entry)

XX DE Mutant human TGF-beta binding protein (BEER) cDNA.

XX KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
XX KW BEER; chromosome 17q12-21; gene therapy; antisense therapy; fracture;
XX KW bone mineralization; mutant; sclerosteosis; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
XX FT CDS 48..119

XX FT /*tag= a

XX FT /note= "mutant BEER created by sclerosteosis
XX FT nonsense mutation"

XX PN WO200032773-A1.

XX PD 08-JUN-2000.

XX PF 24-NOV-1999; 95WO-US27990.

XX PR 27-NOV-1998; 98US-0110283.

XX PA (DARW-) DARWIN DISCOVERY LTD.

XX PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepker BW;

XX PI Van Ness J, Winkler DG;

XX DR WPI; 2000-41321/35.

XX DR P-PSDB; AAY96435.

XX PT Nucleic acids (I) encoding a transforming growth factor beta binding
XX PT protein, useful for identifying agents for treating osteopenia,
XX PT osteoporosis and fractures

XX PS Claim 27; Page 117-118; 162pp; English.

XX CC This cDNA encodes a mutant human transforming growth factor-beta
XX CC (TGF-beta) binding protein (BEER) produced as a result of a nonsense
XX CC mutation in the BEER coding sequence (C to T mutation at position 117).
XX CC This mutation has been linked to sclerosteosis in humans by linkage
XX CC analysis of an Afrikaaner population in South Africa. The hBEER gene has
XX CC been localized to the chromosome 17q12-21. The cDNA and protein may be
XX CC used for prevention, treatment and diagnosis of diseases associated with
XX CC inappropriate BEER expression. For example, they may be used to treat
XX CC disorders associated with decreased TGF-beta BP expression. The cDNA or
XX CC vectors may be administered to treat diseases by rectifying mutations or
XX CC deletions in a patient's genome that affect the activity of BEER by
XX CC expressing inactive proteins or to supplement the patients own production
XX CC of BEER polypeptides. The nucleic acids may be used for recombinant
XX CC production of BEER, gene therapy, antisense therapy, as probes for
XX CC diagnostic assays and for functional studies. BEER may be used to raise
XX CC antibodies and for identification of BEER modulators. BEER antagonists
XX CC may be used to increase bone mineral content for the treatment of
XX CC disorders such as osteopenia, osteoporosis, fractures and other
XX CC disorders associated with low mineral content.

XX SQ Sequence 2301 BP; 568 A; 634 C; 614 G; 485 T; 0 other;

Query Match 10.8%; Score 69; DB 21; Length 2301;

Best Local Similarity 100.0%; Pred. No. 4.2e-23;

Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 475 CTGGTGGCCTCGTGAAGTGCAAGCGCCTCACCCGCTTCCACAAACGAGTGGAGCTCAAG 534

Db 528 CTGGTGGCCTCGTGAAGTGCAAGCGCCTCACCCGCTTCCACAAACGAGTGGAGCTCAAG 587

QY 535 GACTTCGGG 543

Db 588 GACTTCGGG 596

Search completed: March 29, 2003, 08:39:31

Job time : 213.734 secs


```
XX 12-SEP-2000 (first entry)
XX Human TGF-beta binding protein (BEER) cDNA.
XX
XX osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
XX BEER; chromosome 17q12-21; gene therapy; antisense therapy; fracture;
XX bone mineralization; ss.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 48..689
XX FT /*tag= a
XX FT /product= TGF-beta_binding_protein
XX
XX WO200032773-A1.
XX
XX 08-JUN-2000.
XX
XX 24-NOV-1999; 99WO-US27990.
XX
XX 27-NOV-1998; 98US-0110283.
XX
XX (DARW-) DARWIN DISCOVERY LTD.
XX
XX Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepers BW;
XX Van Ness J, Winkler DG;
XX
XX WPI: 2000-412321/35.
XX P-PSDB: AAY96429.
XX
XX Nucleic acids (I) encoding a transforming growth factor beta binding
XX protein, useful for identifying agents for treating osteopenia,
XX osteoporosis and fractures
XX
XX Claim 1; Page 114-115; 162pp; English.
XX
XX This cDNA encodes a human transforming growth factor-beta (TGF-beta)
XX binding protein designated BEER. The hBEER gene has been localized
XX to the chromosome 17q12-21. The cDNA and protein may be used for
XX prevention, treatment and diagnosis of diseases associated with
XX inappropriate BEER expression. For example, they may be used to treat
XX disorders associated with decreased TGF-beta BP expression. The cDNA or
XX vectors may be administered to treat diseases by rectifying mutations or
XX deletions in a patient's genome that affect the activity of BEER by
XX expressing inactive proteins or to supplement the patients own production
XX of BEER polypeptides. The nucleic acids may be used for recombinant
XX production of BEER, gene therapy, antisense therapy, as probes for
XX diagnostic assays and for functional studies. BEER may be used to raise
XX antibodies and for identification of BEER modulators. BEER antagonists
XX may be used to increase bone mineral content for the treatment of
XX disorders such as osteopenia, osteoporosis, fractures and other
XX disorders associated with low mineral content.
XX
XX Sequence 2301 BP; 568 A; 635 C; 614 G; 484 T; 0 other;
XX
XX Query Match 10.8%; Score 69; DB 21; Length 2301;
XX Best Local Similarity 100.0%; Pred. No. 4.2e-23;
XX Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 475 CTGGTGGCTCGTGAAGTGAAGCGCCTCACCGCTTCACACACGATCGGAGCTCAAG 534
DB 528 CTGGTGGCTCGTGAAGTGAAGCGCCTCACCGCTTCACACACGATCGGAGCTCAAG 587
QY 535 GACTTCGGG 543
DB 588 GACTTCGGG 596
RESULT 14
AAA29056
ID AAA29056 standard; cDNA; 2301 BP.
```

```
XX AAA29056;
XX
XX 12-SEP-2000 (first entry)
XX
XX Human TGF-beta binding protein (BEER) variant V10I cDNA.
XX
XX osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
XX BEER; variant; V10I; gene therapy; antisense therapy; fracture;
XX chromosome 17q12-21; bone mineralization; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 48..689
XX FT /*tag= a
XX FT /label= BEER variant V10I
XX FT /product= TGF-beta_binding_protein
XX
XX WO200032773-A1.
XX
XX 08-JUN-2000.
XX
XX 24-NOV-1999; 99WO-US27990.
XX
XX 27-NOV-1998; 98US-0110283.
XX
XX (DARW-) DARWIN DISCOVERY LTD.
XX
XX Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepers BW;
XX Van Ness J, Winkler DG;
XX
XX WPI: 2000-412321/35.
XX P-PSDB: AAY96430.
XX
XX Nucleic acids (I) encoding a transforming growth factor beta binding
XX protein, useful for identifying agents for treating osteopenia,
XX osteoporosis and fractures
XX
XX Claim 1; Page 118-119; 162pp; English.
XX
XX This cDNA encodes a variant human transforming growth factor-beta
XX (TGF-beta) binding protein designated BEER V10I. The encoded protein
XX comprises a substitution of isoleucine for the wild-type valine at
XX residue 10. The cDNA and protein may be used for prevention, treatment
XX and diagnosis of diseases associated with inappropriate BEER expression.
XX For example they may be used to treat disorders associated with
XX decreased TGF-beta BP expression. The cDNA or vectors may be administered
XX to treat diseases by rectifying mutations or deletions in a patient's
XX genome that affect the activity of BEER by expressing inactive proteins
XX or to supplement the patients own production of BEER polypeptides. The
XX nucleic acids may be used for recombinant production of BEER, gene
XX therapy, antisense therapy, as probes for diagnostic assays and for
XX functional studies. BEER may be used to raise antibodies and for
XX identification of BEER modulators. BEER antagonists may be used to
XX increase bone mineral content for the treatment of disorders such as
XX osteopenia, osteoporosis, fractures and other disorders associated with
XX low mineral content.
XX
XX Sequence 2301 BP; 569 A; 634 C; 614 G; 484 T; 0 other;
XX
XX Query Match 10.8%; Score 69; DB 21; Length 2301;
XX Best Local Similarity 100.0%; Pred. No. 4.2e-23;
XX Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 475 CTGGTGGCTCGTGAAGTGAAGCGCCTCACCGCTTCACACACGATCGGAGCTCAAG 534
DB 528 CTGGTGGCTCGTGAAGTGAAGCGCCTCACCGCTTCACACACGATCGGAGCTCAAG 587
QY 535 GACTTCGGG 543
DB 588 GACTTCGGG 596
```



```

CDS      1..642
FT      /*tag= a
FT      /product= "hDCR6 #2"
FT      /partial
XX
XX WO200055193-A2.
XX
XX 21-SEP-2000.
XX
XX 02-MAR-2000; 2000WO-US05537.
XX
XX 12-MAR-1999; 99US-0124118.
XX
XX (REG- ) REGENERON PHARM INC.
XX
XX Economides AN;
XX
XX WPI; 2000-638179/61.
XX P-PSDB; AAB26106.
XX
XX Novel isolated, human DNA/Cerberus related protein 6 which include
XX PT natural homologue, and polypeptides comprising DCR6 domain and nucleic
XX PT acids encoding the proteins which are useful as probes and primers -
XX
XX Claim 2; Fig 3; 40pp; English.
XX
XX The present sequence comprises exons 1 and 4 of the human
XX CC DAN/Cerberus-related protein 6 (hDCR6) coding sequence. It was isolated
XX CC from a human kidney cDNA library containing exons 1 and 4 of the
XX CC sequence. hDCR6 is closely related to the DAN and DCR5 proteins, both of
XX CC which act as antagonists of morphogenic proteins such as BMP. It is
XX CC possible that the hDCR6 gene and protein can be used as immunogens,
XX CC modulators of cell function, growth and differentiation, to reduce
XX CC undesirable bone formation, to identify DCR6 binding agents, in
XX CC diagnosis, and in gene therapy.
XX
XX Sequence 642 BP; 115 A; 230 C; 211 G; 86 T; 0 other;
XX
XX Query Match      10.8%; Score 69; DB 21; Length 642;
XX Best Local Similarity 100.0%; Pred. No. 4.6e-23;
XX Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 475 CTGGTGGCCTCGTGCAAGTGCAGCGCCTCACCGCTTCCACACAGTCGAGCTCAAG 534
XX      |||||||
XX Db 481 CTGGTGGCCTCGTGCAAGTGCAGCGCCTCACCGCTTCCACACAGTCGAGCTCAAG 540
XX
XX QY 535 GACTTCGGG 543
XX      |||||||
XX Db 541 GACTTCGGG 549
XX
XX RESULT 10
XX ABA94293
XX ID ABA94293 standard; cDNA; 759 BP.
XX
XX AC ABA94293;
XX
XX DT 26-MAR-2002 (first entry)
XX
XX DE Human cloaked-2 polypeptide encoding cDNA.
XX
XX KW Cloaked-2; cysteine knot motif; nephrotropic; cardiant; immunomodulator;
XX KW hepatotropic; antiinflammatory; antithyroid; cytostatic; neuroprotective;
XX KW antianemic; hypotensive; antiarrhythmic; antiarteriosclerotic; muscular;
XX KW antidiabetic; anorectic; gene therapy; cell therapy; antisense therapy;
XX KW human; ss.
XX
XX OS Homo sapiens.
XX
XX XX Key Location/Qualifiers
XX FT 37..678
XX CDS /*tag= a
XX FT /product= "cloaked-2 polypeptide"

```

```

FT sig_peptide 37..105
FT FT /*tag= b
FT mat_peptide 106..759
FT FT /*tag= c
FT FT /note= "see ABB07207"
XX
XX WO200192308-A2.
XX
XX 06-DEC-2001.
XX
XX 29-MAY-2001; 2001WO-US17478.
XX
XX 01-JUN-2000; 2000US-208550P.
XX 04-AUG-2000; 2000US-223542P.
XX (AMGE-) AMGEN INC.
XX
XX Paszty CJ, Gao Y;
XX
XX WPI; 2002-114325/15.
XX P-PSDB; ABB07207, ABB07209.
XX
XX New human and mouse cysteine-knot polypeptide designated as Cloaked-2,
XX PT for treating or preventing kidney, heart (e.g. myocardial infarction)
XX PT or liver (e.g. hepatitis) diseases -
XX
XX Claim 1; Fig 1; 170pp; English.
XX
XX The invention relates to polypeptides comprising a cysteine knot motif
XX CC and designated as Cloaked-2, derived from human and mouse. The cloaked-2
XX CC polypeptides can be expressed by standard recombinant methodology. The
XX CC cloaked-2 polynucleotides are useful in gene therapy and antisense
XX CC therapy. The cloaked-2 polypeptides and polynucleotides are useful for
XX CC treating, preventing, ameliorating or detecting diseases and disorders of
XX CC the kidney (e.g. anemia, hypertension or low blood pressure), heart (e.g.
XX CC cardiac hypertrophy, congestive heart failure, myocardial infarction,
XX CC arrhythmias, atherosclerosis, hypertension or low blood pressure),
XX CC skeletal muscle (e.g. muscular dystrophy or cachexia), placenta (e.g.
XX CC congenital abnormalities or miscarriage), liver (e.g. hepatitis or
XX CC cirrhosis), pancreas (e.g. diabetes or pancreatitis), thyroid (e.g.
XX CC Grave's disease or myxedema) or adrenal cortex (e.g. Cushing's disease
XX CC or Addison's disease), homeostasis or metabolic diseases (e.g. obesity,
XX CC cancer or myopathies), infections, or autoimmune diseases. Selective
XX CC binding agents may be used to modulate the biological activities of
XX CC Cloaked-2 polypeptides or to detect Cloaked-2 polypeptide levels in a
XX CC sample. Transgenic non-human animals are useful for drug candidate
XX CC screening. The present sequence represents a cDNA encoding the human
XX CC cloaked-2 polypeptide.
XX
XX Sequence 759 BP; 125 A; 282 C; 244 G; 108 T; 0 other;
XX
XX Query Match      10.8%; Score 69; DB 24; Length 759;
XX Best Local Similarity 100.0%; Pred. No. 4.5e-23;
XX Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 475 CTGGTGGCCTCGTGCAAGTGCAGCGCCTCACCGCTTCCACACAGTCGAGCTCAAG 534
XX      |||||||
XX Db 517 CTGGTGGCCTCGTGCAAGTGCAGCGCCTCACCGCTTCCACACAGTCGAGCTCAAG 576
XX
XX QY 535 GACTTCGGG 543
XX      |||||||
XX Db 577 GACTTCGGG 585
XX
XX RESULT 11
XX AAA94050
XX ID AAA94050 standard; DNA; 1104 BP.
XX
XX AC AAA94050;
XX
XX DT 15-JAN-2001 (first entry)
XX
XX DE Human DAN/Cerberus-related protein 6 (hDCR6) coding sequence #1.

```

PT analyzing gene expression in human bone marrow -

XX Example 4; SEQ ID NO: 20066; 658pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
 CC the probes of the invention.

XX Sequence 422 BP; 53 A; 145 C; 159 G; 65 T; 0 other;

Query Match 10.8%; Score 69; DB 23; Length 422;

Best Local Similarity 100.0%; Pred. No. 4.7e-23;

Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 475 CTGGTGGCTCTGTCAGTGCAGCGCTTCCACCGCTTCCACACAGTGGAGCTCAAG 534

Db 162 CTGGTGGCTCTGTCAGTGCAGCGCTTCCACCGCTTCCACACAGTGGAGCTCAAG 103

Qy 535 GACTTCGGG 543

Db 102 GACTTCGGG 94

RESULT 8

ABS19777/C

ID ABS19777 standard; DNA; 422 BP.

XX ABS19777;

XX ABS19777;

XX 19-AUG-2002 (first entry)

XX Human genome-derived single exon probe ORF from lung SEQ ID No 19768.

XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease; open reading frame; ORF.

XX Homo sapiens.

XX WO200186003-A2.

XX 15-NOV-2001.

XX 30-JAN-2001; 2001WO-US00665.

XX 04-FEB-2000; 2000US-180312P.

XX 26-MAY-2000; 2000US-207456P.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-234687P.

XX 27-SEP-2000; 2000US-236359P.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2002-114183/15.

XX Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -

XX Claim 4; SEQ ID No 19768; 634pp; English.

XX

CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridise at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a single exon
 CC probe open reading frame of the invention.

CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/publ:shed_pct_sequences.

XX Sequence 422 BP; 53 A; 145 C; 159 G; 65 T; 0 other;

Query Match 10.8%; Score 69; DB 24; Length 422;

Best Local Similarity 100.0%; Pred. No. 4.7e-23;

Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 475 CTGGTGGCTCTGTCAGTGCAGCGCTTCCACCGCTTCCACACAGTGGAGCTCAAG 534

Db 162 CTGGTGGCTCTGTCAGTGCAGCGCTTCCACCGCTTCCACACAGTGGAGCTCAAG 103

Qy 535 GACTTCGGG 543

Db 102 GACTTCGGG 94

RESULT 9

AAA94051

ID AAA94051 standard; DNA; 642 BP.

XX AAA94051;

XX 15-JAN-2001 (first entry)

XX Human DAN/Cerberus-related protein 6 (hDCR6) cDNA exons 1 and 4.

XX Human; DNA/Cerberus-related protein 6; hDCR6; morphogenic protein;
 KW antagonist; BMP; cell growth; cell differentiation; bone formation;
 KW gene therapy; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

Db 280 CTGGTGGCCTGCAAGTGCAGCGCCTCACCCTCCACACAGTCGAGCTCAAG 221
 QY 535 GACTTCGGG 543
 Db 220 GACTTCGGG 212
 RESULT 6
 ID ABS07272/c
 ID ABS07272 standard; DNA; 392 BP.
 AC ABS07272;
 XX
 XX 19-AUG-2002 (first entry)
 XX Human genome-derived single exon probe from lung SEQ ID No 7263.
 DE Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberosus sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioma; lymphoma; Karagener syndrome;
 KW primary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 XX Homo sapiens.
 OS
 OS WO200186003-A2.
 PN
 XX
 XX 15-NOV-2001.
 XX
 XX 30-JAN-2001; 2001WO-US00665.
 XX
 PR 04-FEB-2000; 2000US-180312P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236359P.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI
 XX WPI; 2002-114183/15.
 DR
 XX Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 XX
 XX Claim 1; SEQ ID No 7263; 634pp; English.
 XX
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridise at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several

CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberosus sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosiderosis, pulmonary histiocytosis, lymphangioma, lymphoma,
 CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a single exon
 CC probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 392 BP; 60 A; 123 C; 149 G; 60 T; 0 other;
 Query Match 10.8%; Score 69; DB 24; Length 392;
 Best Local Similarity 100.0%; Pred. No. 4.7e-23;
 Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 475 CTGGTGGCCTGCAAGTGCAGCGCCTCACCCTCCACACAGTCGAGCTCAAG 534
 Db 280 CTGGTGGCCTGCAAGTGCAGCGCCTCACCCTCCACACAGTCGAGCTCAAG 221
 QY 535 GACTTCGGG 543
 Db 220 GACTTCGGG 212
 RESULT 7
 AAK45509/c
 ID AAK45509 standard; DNA; 422 BP.
 XX
 AC AAK45509;
 XX
 XX 06-NOV-2001 (first entry)
 DT Human bone marrow expressed single exon probe SEQ ID NO: 20066.
 DE Human; bone marrow expressed exon; gene expression analysis; probe;
 KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
 XX Homo sapiens.
 OS
 XX WO200157276-A2.
 PN
 XX 09-AUG-2001.
 PD
 XX 30-JAN-2001; 2001WO-US00668.
 PF
 XX 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 FA
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI
 XX WPI; 2001-488900/53.
 DR
 XX Human genome-derived single exon nucleic acid probes useful for

CC associated with low mineral content.

Sequence 638 BP; 114 A; 224 C; 209 G; 91 T; 0 other;
XX
SQ

Query Match 92.0%; Score 585; DB 21; Length 638;
Best Local Similarity 99.8%; Pred. NO. 2.3e-269;
Matches 635; Conservative 0; Mismatches 1; Indels

Qy	1	ATGCAGCGCTCACTAGCGCCGTGCTCATCTGCGCTATTGTGCACGTCGCTTCTGTGCT	60
Db	1	ATGCAGCGCTCACTAGCGCCGTGCTCATCTGCGCTATTGTGCACGTCGCTTCTGTGCT	60
Qy	61	GTGAGGCGCCAGGCGTGGCAAGCCTTCAGGAATGATGCCACAGAGGTCATCCACAGGCTT	120
Db	61	GTGAGGGCCAGGGGTGGCAAGCCTTCAGGAATGATGCCACAGAGTCATCCACAGGCTT	120
Qy	121	GGAGAGTACCCCGAGCCTCTCTTGAGAACAACAGACCATGAACCGGGCGGAGATGGA	180
Db	121	GGAGAGTACCCCGAGCCTCTCTTGAGAACAACAGACCATGAACCGGGCGGAGATGGA	180
Qy	181	GGCAGACCTCCCAACCATCCTATGAGCGCCAAAGATGTGTCCAGATACAGTCCGCGAG	240
Db	181	GGCAGACCTCCCAACCATCCTATGAGCGCCAAAGATGTGTCCAGATACAGTCCGCGAG	240
Qy	241	CTGCACATACACCCGCTTCCTGACAGACGGCCCATGCGCAGCGCCAAAGCCGGTCACCGAG	300
Db	241	CTGCACATACACCCGCTTCCTGACAGACGGCCCATGCGCAGCGCCAAAGCCGGTCACCGAG	300
Qy	301	TTGTGTGTCTCCGCGCCAGTGGGGCCCCGCGCGGTGTGTCGCCAACGCCATCGGGCGCGTG	360
Db	301	TTGTGTGTCTCCGCGCCAGTGGGGCCCCGCGCGGTGTGTCGCCAACGCCATCGGGCGCGTG	360
Qy	361	AAGTGTGGCGCCGACGAGCAGCATTCCTCGCTGCATCCCGATCGCTACCGCGCGAG	420
Db	361	AAGTGTGGCGCCGACGAGCAGCATTCCTCGCTGCATCCCGATCGCTACCGCGCGAG	420
Qy	421	CGGTGTCAGCTGTGTGCTCCCGGGGGCGCGCGCTCGCGCAAGGTGCGTCTGTGGT	480
Db	421	CGGTGTCAGCTGTGTGCTCCCGGGGGCGCGCGCTCGCGCAAGGTGCGTCTGTGGT	480
Qy	481	GCCTGTGTGAAGTGCAAGCGCTCAACCGCTTCCAAACAGTTCGAGCTCAAGGACTTC	540
Db	481	GCCTGTGTGAAGTGCAAGCGCTCAACCGCTTCCAAACAGTTCGAGCTCAAGGACTTC	540
Qy	541	GGGCGGAGACCGCGCGCGCGCAGAGGGTGCAGCGCGGCCCGCGCCCGGGGAGCC	600
Db	541	GGGCGGAGACCGCGCGCGCGCAGAGGGTGCAGCGCGGCCCGCGCCCGGGGAGCC	600
Qy	601	AAAGCCAAACAGCGGAGCTGGAGAACGCTACTAG	636
Db	601	AAAGCCAAACAGCGGAGCTGGAGAACGCTACTAG	636

RESULT 3

AAA29063
ID AAA29063 standard; DNA; 35828 BP.

XX
AC

XX
DT 12-SEP-2000 (first entry)

XX Murine TGF-beta binding protein (BBER) genomic MluI-AviIII DNA fragment.
DE
XX
XX Osteopathic; transforming growth factor-beta, TGF-beta; binding protein
KW BBER; gene therapy; antineoplastic therapy; fracture; bone mineralization;
KW transgenic mouse; over-expression; ds.

XX
OS Mus musculus.

XX		
FH	Key	Location/Qualifiers
FT	CDS	18317..21464
FT		/tag= a
FT	exon	18317..18530

FT		/tag= b
FT	intron	/partial
FT		18531..21022
FT		/tag= c
FT	exon	21023..21464
FT		/tag= d
FT		/partial
XX		
XX		
BN	WO200032773-A1.	
XX		
PD	08-JUN-2000.	
XX		
PF	24-NOV-1999;	99WO-US27990.
XX		
PR	27-NOV-1998;	98US-0110283.
XX		
PA	(DARW-) DARWIN DISCOVERY LTD	
XX		
PI	Brunkow ME, Galas DJ, Kovacs	
PI	Van Nessel J, Winkler DG;	
PI		
DR	WPI; 2000-412321/35.	
DR	P-PSDB; AAY96432.	
DR		

PT Nucleic acids (I) encoding a transforming growth factor beta binding
PT protein, useful for identifying agents for treating osteopenia,
PT osteoporosis and fractures

PS Example 9; Page 74-87; 162pp; English.

This 35 kb MLUI-AvIII genomic DNA restriction fragment includes the entire murine transforming growth factor-beta (TGF-beta) binding protein (BEER) gene. This DNA was used to generate a transgenic mouse over-expressing the BEER gene. The gene and protein may be used for prevention, treatment and diagnosis of diseases associated with inappropriate BEER expression. For example, they may be used to treat disorders associated with decreased TGF-beta BP expression. The cDNA or vectors may be administered to treat diseases by rectifying mutations or deletions in a patient's genome that affect the activity of BEER by expressing inactive proteins or to supplement the patients own production of BEER polypeptides. The nucleic acids may be used for recombinant production of BEER, gene therapy, antisense therapy, as probes for diagnostic assays and for functional studies. BEER may be used to raise antibodies and for identification of BEER modulators. BEER antagonists may be used to increase bone mineral content for the treatment of disorders such as osteopenia, osteoporosis, fractures and other disorders associated with low mineral content.

Sequence 35828 BP; 8973 A; 8817 C; 9064 G; 8973 T; 1 other;

Query Match	66.7%	Score 424;	DB 21;	Length 35828;
Best Local Similarity	100.0%;	Pred. No. 1.2e-192;		
Matches 424;	Conservative	0;	Mismatches 0;	Indels 0;
Gaps				0;

Qy	213	AGATGTTCCGAGTACAGCTGCCGCGAGCTGCACACACCGCTTCTGTACAGACGGCCC	272
Db	21021	AGATGTTCCGAGTACAGCTGCCGCGAGCTGCACACACCGCTTCTGTACAGACGGCCC	21080
Qy	273	ATGCCGAGCGCCAAAGCCGCTACCCAGTTGGTGTCTCGGCCAGTGCGGCCCGCGCG	332
Db	21081	ATGCCGAGCGCCAAAGCCGCTACCCAGTTGGTGTCTCGGCCAGTGCGGCCCGCGCG	21140
Qy	333	GTTGCTGCCAAAGCCCATCGGGCGCGTGAAGTGGTGGCGCCCGAACCGGATTTCCG	392
Db	21141	GCTGCTCCCAAAGCCCATCGGGCGCGTGAAGTGGTGGCGCCCGAACCGGATTTCCG	21200
Qy	393	CTGCATCCCGGATCGCTACCGCGCGACGCGGTGCAGCTGTGTGCCCGGGGGCGGGC	452
Db	21201	CTGCATCCCGATCGCTACCGCGCGACGCGGTGCAGCTGTGTGCCCGGGGGCGGGC	21260
Qy	453	GCGCGCTCGCGAAAGTGCGTCTGTGGCTCTGTCAAGTGCAAGCGCTTCACCCCGTT	512
Db	21261	GCGCGCTCGCGAAAGTGCGTCTGTGGCTCTGTCAAGTGCAAGCGCTTCACCCCGTT	21320

PF 29-MAY-2001; 2001WO-US17478.
XX
PR 01-JUN-2000; 2000US-208550P.
PR 04-AUG-2000; 2000US-223542P.
XX
XX
PA (AMGE-) AMGEN INC.
PI Paszty CJ, Gao Y;
XX
XX WPI: 2002-114325/15.
DR P-PSDB; ABB07208, ABB07210.
XX
XX
PT New human and mouse cysteine-knot polypeptide designated as Cloaked-2,
PT for treating or preventing kidney, heart (e.g. myocardial infarction)
PT or liver (e.g. hepatitis) diseases
XX
XX Claim 1; Fig 2; 170pp; English.
XX
CC The invention relates to polypeptides comprising a cysteine knot motif
CC and designated as Cloaked-2, derived from human and mouse. The cloaked-2
CC polypeptides can be expressed by standard recombinant methodology. The
CC cloaked-2 polynucleotides are useful in gene therapy and antisense
CC therapy. The cloaked-2 polypeptides and polynucleotides are useful for
CC treating, preventing, ameliorating or detecting diseases and disorders of
CC the kidney (e.g. anemia, hypertension or low blood pressure), heart (e.g.
CC cardiac hypertrophy, congestive heart failure, myocardial infarction,
CC arrhythmias, atherosclerosis, hypertension or low blood pressure),
CC skeletal muscle (e.g. muscular dystrophy or cachexia), placenta (e.g.
CC congenital abnormalities or miscarriage), liver (e.g. hepatitis or
CC cirrhosis), pancreas (e.g. diabetes or pancreatitis), thyroid (e.g.
CC Grave's disease or myxedema) or adrenal cortex (e.g. Cushing's disease
CC or Addison's disease), homeostasis or metabolic diseases (e.g. obesity,
CC cancer or myopathies), infections, or autoimmune diseases. Selective
CC binding agents may be used to modulate the biological activities of
CC Cloaked-2 polypeptides or to detect Cloaked-2 polypeptide levels in a
CC sample. Transgenic non-human animals are useful for drug candidate
CC screening. The present sequence represents a cDNA encoding the mouse
CC cloaked-2 polypeptide.
XX
SQ Sequence 636 BP; 114 A; 224 C; 207 G; 91 T; 0 other;

Query Match 100.0%; Score 636; DB 24; Length 636;
Best Local Similarity 100.0%; Pred. No. 1e-293;
Matches 636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCAGCCCTCACTAGCCCGGCTCATCTGCTACTTGTGACGCTGCTCTGCTCT 60
Db |||||
QY 61 GTGGAGGGCCAGGGTGGCAAGCCTTCAGGAATGATGCCACAGAGGTATCCAGGGCTT 120
Db |||||
QY 121 GGAGAGTACCCGAGCTCTCTCTGAGAACACACAGACCATGAAACCGGGGAGAAATGGA 180
Db |||||
QY 181 GGCAGACCTCCACCATCTATGACGCGCAAGATGTTCGAGTACAGCTGCCGCGAG 240
Db |||||
QY 241 CTGCACTACACCGCTTCTTGACAGACGCGCCCATGCCGAGCGCAACCGGTACCCGAG 300
Db |||||
QY 301 TTGGTGTGCTCGGCGAGTGCAGGCGCGGCTGTGCTGCCAACGCCATCGGGGCGGTG 360
Db |||||
QY 361 AAGTGGTGGCGCCGAGACCGGATTTCCGCTGCATCCCGGATTCGTTACCGGCGGAG 420
Db |||||

QY 421 CGGGTGCAGCTGTGTGCCCGGGGGCGCGCGCTCGCAAGTGCCTGTGGTG 480
Db |||||
QY 421 CGGGTGCAGCTGTGTGCCCGGGGGCGCGCGCTCGCAAGTGCCTGTGGTG 480
Db |||||
QY 481 GCCTGTGCAAGTGCAGCGCTTCCACCAACAGTCCGAGCTCAAGGACTTC 540
Db |||||
QY 541 GGGCGGAGACCGCGCGCGCGAGAGGTCGCAAGCGCGCGCGCGCGCGGAGCC 600
Db |||||
QY 601 AAAGCCCAACAGCGGAGCTGGAGAACGCTTACTAG 636
Db |||||
QY 601 AAAGCCCAACAGCGGAGCTGGAGAACGCTTACTAG 636
Db |||||

RESULT 2

AAA29058
ID AAA29058 standard; cDNA; 638 BP.
XX
AC AAA29058;
XX
DT 12-SEP-2000 (first entry)
XX
DE Murine TGF-beta binding protein (BEER) cDNA.
XX
KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
KW BEER; gene therapy; antisense therapy; fracture; bone mineralization; ss.
XX
OS Mus musculus.
PH Key Location/Qualifiers
FT CDS 1..636
FT /*tag= a
FT /product= TGF-beta_binding_protein
XX
PN WO200032773-A1.

PD 08-JUN-2000.

XX 24-NOV-1999; 99WO-US27990.

XX 27-NOV-1998; 98US-0:10283.

XX (DARW-) DARWIN DISCOVERY LTD.

XX Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepers BW;
PI Van Ness J, Winkler DG;
XX WPI: 2000-412321/35.
DR P-PSDB; AAY96432.

XX Nucleic acids (I) encoding a transforming growth factor beta binding
PT protein, useful for identifying agents for treating osteopenia,
PT osteoporosis and fractures
XX

PS Claim 1; Page 123; 16pp; English.

XX This cDNA encodes a murine transforming growth factor-beta (TGF-beta)
CC binding protein designated mBEER. The cDNA and protein may be used for
CC prevention, treatment and diagnosis of diseases associated with
CC inappropriate BEER expression. For example, they may be used to treat
CC disorders associated with decreased TGF-beta BP expression. The cDNA or
CC vectors may be administered to treat diseases by rectifying mutations or
CC deletions in a patient's genome that affect the activity of BEER by
CC expressing inactive proteins or to supplement the patients own production
CC of BEER polypeptides. The nucleic acids may be used for recombinant
CC production of BEER, gene therapy, antisense therapy, as probes for
CC diagnostic assays and for functional studies. BEER may be used to raise
CC antibodies and for identification of BEER modulators. BEER antagonists
CC may be used to increase bone mineral content for the treatment of
CC disorders such as osteopenia, osteoporosis, fractures and other disorders

Mon Mar 31 09:44:28 2003

```

VERSION      AF326739.1  GI:13161019
KEYWORDS
SOURCE
ORGANISM      Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS      Brunkow,M.E., Gardner,J.C., Van Ness,J., Paepker,B.W.,
Kovacevich,B.R., Proll,S., Skonier,J.E., Zhao,L., Sabo,P.J.,
Fu,Y.H., Alisch,R.S., Gillett,L., Colbert,T., Tacconi,P., Galas,D.,
Hamersma,H., Beighton,P. and Mulligan,J.T.
TITLE      Bone dysplasia sclerostosis results from loss of the SOST gene
product, a novel cysteine knot-containing protein
JOURNAL      Am. J. Hum. Genet. 68 (3), 577-589 (2001)
MEDLINE      21090529
PUBMED      11179006
REFERENCE
AUTHORS      Brunkow,M.E., Gardner,J.C., Van Ness,J., Paepker,B.W.,
Kovacevich,B.R., Proll,S., Skonier,J.E., Zhao,L., Sabo,P.J.,
Fu,Y.H., Alisch,R.S., Gillett,L., Colbert,T., Tacconi,P., Galas,D.,
Hamersma,H., Beighton,P. and Mulligan,J.T.
TITLE      Direct Submission
JOURNAL      Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631
220th St. SE, Bothell, WA 98021, USA
FEATURES
source
1. .2323
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /chromosome="17"
   /map="17q21"
1. .47
48. .689
   /codon_start=1
   /product="sclerostin"
   /protein_id="AAK13454.1"
   /db_xref="GI:13161020"
   /translation="MOLPLALCIIVLHVHTAPRVVEGQWQAFKNDATETIIPELGEY
EPPELENNKTNRAENGRRPHHPFETKDVSEYSCRELHFTYVYTDGPCRSAPVTE
LVCSGQCGPARLLPNAIGKWKWRPSGDPFCIPDRYRAORVQLLPQCGEAPRARKVR
LVASCKCKRLTRFHNSLKFQGTAAARFQKGRKPRFRARSAKANQAELENAY"
3'UTR      690..2323
BASE COUNT  590 a 635 c 614 g 484 t
ORIGIN
Query Match      10.8%; Score 69; DB 9; Length 2323;
Best Local Similarity 100.0%; Pred. No. 3.6e-23;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 475 CTGGTGGCCTCGTCAAGTCAAGCGCTCACCGCTTCCACCAACCACTCGGAGCTCAAG 534
Db 528 CTGGTGGCCTCGTCAAGTCAAGCGCTCACCGCTTCCACCAACCACTCGGAGCTCAAG 587
Qy 535 GACTTCGGG 543
Db 588 GACTTCGGG 596
Search completed: March 29, 2003, 10:01:47
Job time : 2363.75 secs

```

```

Db 42072 CGCTTCCACACAGTCGGAGCTCAAGGACTTCGG 42106
|||||
RESULT 12
AX323453 759 bp DNA linear PAT 07-JAN-2002
LOCUS
DEFINITION Sequence 1 from Patent WO0192308.
ACCESSION AX323453
VERSION AX323453.1 GI:18094216
KEYWORDS human.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Passty,C.J. and Gao,Y.
TITLE Cystine-knot polypeptides: cloaked-2 molecules and uses thereof
JOURNAL Patent: WO 0192308-A 1 06-DEC-2001;
Amgen, Inc. (US)
FEATURES
location/Qualifiers
source
BASE COUNT 125 a 282 c 244 g 108 t
ORIGIN
Query Match 10.8%; Score 69; DB 6; Length 759;
Best Local Similarity 100.0%; Pred. No. 4.3e-23;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 475 CTGGTGGCTCTGCAAGTGAAGCGCTTCCACACAGTCGGAGCTCAAG 534
|||||
Db 517 CTGGTGGCTCTGCAAGTGAAGCGCTTCCACACAGTCGGAGCTCAAG 576
|||||
Qy 535 GACTTCGGG 543
|||||
Db 577 GACTTCGGG 585
|||||

RESULT 13
AX342535 2271 bp DNA linear PAT 12-JAN-2002
LOCUS
DEFINITION Sequence 1 from Patent WO0198491.
ACCESSION AX342535
VERSION AX342535.1 GI:18151964
KEYWORDS human.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Bailemans,W., Ebeling,M., Foernzler,D., Patel,N., van Hul,W. and
Vickery,B.H.
TITLE Osteolevin gene polymorphisms
JOURNAL Patent: WO 0198491-A 1 27-DEC-2001;
F. HOFFMANN-LA ROCHE AG (CH); UNIVERSITAIRE INSTELLING ANTWERPEN
(BE)
FEATURES
location/Qualifiers
source
BASE COUNT 573 a 623 c 599 g 476 t
ORIGIN
Query Match 10.8%; Score 69; DB 6; Length 2271;
Best Local Similarity 100.0%; Pred. No. 3.6e-23;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 475 CTGGTGGCTCTGCAAGTGAAGCGCTTCCACACAGTCGGAGCTCAAG 534
|||||
Db 493 CTGGTGGCTCTGCAAGTGAAGCGCTTCCACACAGTCGGAGCTCAAG 552
|||||

Qy 535 GACTTCGGG 543
|||||
Db 553 GACTTCGGG 561
|||||

RESULT 14
AF331844 2296 bp mRNA linear PRI 06-MAR-2001
LOCUS
DEFINITION Homo sapiens SOST (SOST) mRNA, complete cds.
ACCESSION AF331844
VERSION AF331844.1 GI:13236417
KEYWORDS
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 2296)
Bailemans,W., Ebeling,M., Patel,N., van Hul,W., Olson,P.,
Dioszegi,M., Lacza,C., Wuyts,W., van den Ende,J., Willems,P.,
Paes-Alves,A.F., Hill,S., Bueno,M., Ramos,F.J., Tacconi,P.,
Dijkers,F.G., Stratakis,C., Lindpaintner,K., Vickery,B.,
Foernzler,D. and Van Hul,W.
TITLE Increased bone density in sclerosteosis is due to the deficiency of
a novel secreted protein (SOST)
JOURNAL Hum. Mol. Genet. 10 (5), 537-543 (2001)
MEDLINE 21096930
PUBMED 11181578
REFERENCE
AUTHORS 2 (bases 1 to 2296)
Bailemans,W., Ebeling,M., Patel,N., Vickery,B., Foernzler,D. and Van
Hul,W.
DIRECT SUBMISSION
TITLE Submitted (22-DEC-2000) Medical Genetics, University of Antwerp,
Universiteitssp.lein 1, Antwerp B2610, Belgium
JOURNAL
FEATURES
location/Qualifiers
source
BASE COUNT 576 a 631 c 607 g 482 t
ORIGIN
Query Match 10.8%; Score 69; DB 9; Length 2296;
Best Local Similarity 100.0%; Pred. No. 3.6e-23;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 475 CTGGTGGCTCTGCAAGTCGCAAGCGCTTCCACACAGTCGGAGCTCAAG 534
|||||
Db 518 CTGGTGGCTCTGTCGCAAGTCGCAAGCGCTTCCACACAGTCGGAGCTCAAG 577
|||||
Qy 535 GACTTCGGG 543
|||||
Db 578 GACTTCGGG 586
|||||

RESULT 15
AF326739 2323 bp mRNA linear PRI 28-FEB-2001
LOCUS
DEFINITION Homo sapiens sclerostin mRNA, complete cds.
ACCESSION AF326739

```

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

Project Information
Center project name: GRII
Center clone name: CH230-137114

Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 49288 bases at least Q40
Consensus quality: 51947 bases at least Q30
Consensus quality: 54106 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 54 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 1021: contig of 1021 bp in length
* 1022 1121: gap of unknown length
* 1122 2201: contig of 1080 bp in length
* 2202 2301: gap of unknown length
* 2302 3754: contig of 1453 bp in length
* 3755 3854: gap of unknown length
* 3855 5142: contig of 1288 bp in length
* 5143 5242: gap of unknown length
* 5243 6277: contig of 1035 bp in length
* 6278 6377: gap of unknown length
* 6378 7608: contig of 1231 bp in length
* 7609 7708: gap of unknown length
* 7709 8737: contig of 1029 bp in length
* 8738 8837: gap of unknown length
* 8838 10296: contig of 1459 bp in length
* 10297 10396: gap of unknown length
* 10397 11476: contig of 1080 bp in length
* 11477 11576: gap of unknown length
* 11577 12612: contig of 1036 bp in length
* 12613 12712: gap of unknown length
* 12713 13734: contig of 1022 bp in length
* 13735 13834: gap of unknown length
* 13835 14980: contig of 1146 bp in length
* 14981 15080: gap of unknown length
* 15081 16334: contig of 1254 bp in length
* 16335 16434: gap of unknown length
* 16435 17578: contig of 1144 bp in length
* 17579 17678: gap of unknown length
* 17679 19476: contig of 1798 bp in length
* 19477 19576: gap of unknown length
* 19577 20675: contig of 1099 bp in length
* 20676 20775: gap of unknown length
* 20776 22480: contig of 1705 bp in length
* 22481 22580: gap of unknown length
* 22581 23838: contig of 1258 bp in length
* 23839 23938: gap of unknown length
* 23939 25468: contig of 1530 bp in length
* 25469 25568: gap of unknown length
* 25569 27408: contig of 1840 bp in length
* 27409 27508: gap of unknown length
* 27509 28800: contig of 1292 bp in length
* 28801 28900: gap of unknown length
* 28901 2942: contig of 1042 bp in length
* 2943 30042: gap of unknown length
* 30043 31521: contig of 1479 bp in length
* 31522 31621: gap of unknown length
* 31622 33255: contig of 1634 bp in length
* 33256 33255: gap of unknown length

* 33356 34559: contig of 1304 bp in length
* 34559 34759: gap of unknown length
* 34760 36076: contig of 1317 bp in length
* 36077 36176: gap of unknown length
* 36177 37895: contig of 1719 bp in length
* 37896 37995: gap of unknown length
* 37996 39735: contig of 1740 bp in length
* 39736 39835: gap of unknown length
* 39836 41383: contig of 1548 bp in length
* 41384 41484: gap of unknown length
* 41484 43588: contig of 2105 bp in length
* 43589 43688: gap of unknown length
* 43689 45634: contig of 1945 bp in length
* 45634 45734: gap of unknown length
* 45734 47530: contig of 1797 bp in length
* 47531 47630: gap of unknown length
* 47631 48953: contig of 1323 bp in length
* 48954 49053: gap of unknown length
* 49054 51099: contig of 2046 bp in length
* 51100 51999: gap of unknown length
* 52829: contig of 1630 bp in length
* 52830 52929: gap of unknown length
* 52930 55451: contig of 2522 bp in length
* 55452 55551: gap of unknown length
* 55552 58096: contig of 2545 bp in length
* 58097 58196: gap of unknown length
* 58197 60127: contig of 1931 bp in length
* 60128 62306: contig of 2079 bp in length
* 62307 62406: gap of unknown length
* 62407 64228: contig of 1822 bp in length
* 64229 64328: gap of unknown length
* 64329 66422: contig of 2094 bp in length
* 66423 66522: gap of unknown length
* 66523 68670: contig of 2148 bp in length
* 68671 68770: gap of unknown length
* 68771 71676: contig of 2906 bp in length
* 71677 71776: gap of unknown length
* 71777 74664: contig of 2888 bp in length
* 74665 74764: gap of unknown length
* 74765 77233: contig of 2469 bp in length
* 77234 77333: gap of unknown length
* 77334 78578: contig of 1245 bp in length
* 78579 78678: gap of unknown length
* 78679 81968: contig of 3290 bp in length
* 81969 82068: gap of unknown length
* 82069 84118: contig of 2350 bp in length
* 84119 84518: gap of unknown length
* 84519 88073: contig of 3555 bp in length
* 88074 88173: gap of unknown length
* 88174 91080: contig of 2907 bp in length
* 91081 91180: gap of unknown length
* 91181 93633: contig of 2453 bp in length
* 93634 93733: gap of unknown length
* 93734 97324: contig of 3591 bp in length
* 97325 97424: gap of unknown length
* 97425 101506: contig of 4082 bp in length
* 101507 101606: gap of unknown length
* 101607 104898: contig of 3292 bp in length.

Query Match 16.4%; Score 104; DB 2; Length 104898;
Best Local Similarity 99.4%; Pred. No. 8.5e-41;
Matches 154; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 388 TTCCGCTGCATCCCGGATCGTACCGCGCAGCGGTGCTGTCGCCCGGGGGC 447
Db 41952 TTCCGCTGCATCCCGGATCGTACCGCGCAGCGGTGCTGTCGCCCGGGGGC 42011
Qy 448 CGCGCGCGCGCTCGCGCAAGGTGCGTCTGGTGCCTCGTCAAGTGAAGCGCCTCACC 507
Db 42012 CGCGCGCGCGCTCGCGCAAGGTGCGTCTGGTGCCTCGTCAAGTGAAGCGCCTCACC 42071
Qy 508 CGCTTCCACACACAGTCCGAGCTCAAGACTTCGG 542

RESULT 10
AC098160/c

LOCUS
DEFINITION

AC098160 101804 bp DNA linear HTG 12-JUL-2002
Rattus norvegicus clone CH230-60M11, *** SEQUENCING IN PROGRESS
***, 53 unordered pieces.

ACCESSION
VERSION

AC098160.3 GI:21729002

KEYWORDS

HTG; HTGS PHASE1.

SOURCE
ORGANISM

Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus

REFERENCE
AUTHORS

1 (bases 1 to 101804)
Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C.,
Albrooks, S.L., Amarantunge, H.C., Are, J.R., Ayale, M., Banks, T.,
Barbata, J., Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D.,
Bouck, J., Bowie, S., Brivea, M., Brown, E., Brown, M., Bryant, N.P.,
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
DeLaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
Hernandez, O., Hodgson, A., Hognes, M., Holloway, C., Hollins, B.,
Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Joudan, S.,
Jacobson, B., Jia, Y., Johnson, J., Jolivet, S., Joudan, S.,
Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseg, H.,
Lozano, R., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,
Massey, E., McWhiney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M.,
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokweto, S., Ogih, M., Okwuonu, G.,
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
Rives, M., Rojas, A., Rojibokan, I., Roife, M., Ruiz, S., Savary, G.,
Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I.,
Sodergren, E., Sonaite, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usmani, K., Vazquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 101804)
Worley, K.C.
Direct Submission
Submitted (23-OCT-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 101804)
Worley, K.C.
Direct Submission
Submitted (12-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 11, 2002 this sequence version replaced gi:17972960.
----- Genome Center
Center: Baylor College of Medicine

TITLE
JOURNAL

REFERENCE
AUTHORS

1 (bases 1 to 101804)
Worley, K.C.
Direct Submission
Submitted (23-OCT-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 101804)
Worley, K.C.
Direct Submission
Submitted (12-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 11, 2002 this sequence version replaced gi:17972960.
----- Genome Center
Center: Baylor College of Medicine

REFERENCE
AUTHORS

1 (bases 1 to 101804)
Worley, K.C.
Direct Submission
Submitted (23-OCT-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 101804)
Worley, K.C.
Direct Submission
Submitted (12-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 11, 2002 this sequence version replaced gi:17972960.
----- Genome Center
Center: Baylor College of Medicine

TITLE
JOURNAL

1 (bases 1 to 101804)
Worley, K.C.
Direct Submission
Submitted (23-OCT-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 101804)
Worley, K.C.
Direct Submission
Submitted (12-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 11, 2002 this sequence version replaced gi:17972960.
----- Genome Center
Center: Baylor College of Medicine

COMMENT

1 (bases 1 to 101804)
Worley, K.C.
Direct Submission
Submitted (23-OCT-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 101804)
Worley, K.C.
Direct Submission
Submitted (12-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 11, 2002 this sequence version replaced gi:17972960.
----- Genome Center
Center: Baylor College of Medicine

Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GGVF
Center clone name: CH230-60M11
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 42516 bases at least Q40
Consensus quality: 46513 bases at least Q30
Consensus quality: 49790 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 53 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1
* 1001: contig of 1001 bp in length
* 1101: gap of unknown length
* 1102: contig of 1036 bp in length
* 2138: gap of unknown length
* 2238: contig of 1380 bp in length
* 3618: gap of unknown length
* 4928: contig of 1210 bp in length
* 5027: gap of unknown length
* 5028: contig of 1482 bp in length
* 6510: gap of unknown length
* 8256: contig of 1647 bp in length
* 8357: gap of unknown length
* 8577: contig of 1351 bp in length
* 9708: gap of unknown length
* 11221: contig of 1414 bp in length
* 11222: gap of unknown length
* 11322: contig of 1459 bp in length
* 12781: gap of unknown length
* 12880: gap of unknown length
* 13909: contig of 1029 bp in length
* 14009: gap of unknown length
* 15194: contig of 1185 bp in length
* 15294: gap of unknown length
* 15295: contig of 1164 bp in length
* 16458: gap of unknown length
* 16559: gap of unknown length
* 17876: contig of 1318 bp in length
* 17976: gap of unknown length
* 19154: contig of 1178 bp in length
* 19254: gap of unknown length
* 20969: contig of 1715 bp in length
* 21089: gap of unknown length
* 22933: contig of 1864 bp in length
* 23033: gap of unknown length
* 24220: contig of 1187 bp in length
* 24320: gap of unknown length
* 26128: contig of 1808 bp in length
* 26228: gap of unknown length
* 27629: contig of 1467 bp in length
* 27795: gap of unknown length
* 29044: contig of 1249 bp in length
* 29144: gap of unknown length
* 30201: contig of 1057 bp in length
* 30301: gap of unknown length
* 31925: contig of 1624 bp in length
* 32025: gap of unknown length
* 33523: contig of 1498 bp in length
* 33623: gap of unknown length
* 34759: contig of 1136 bp in length
* 34859: gap of unknown length
* 36227: contig of 1368 bp in length
* 34860

Db 60 CTCACCCGCTTCCACCAAGTCGAGCTCAAGAGCTTGAGACTGAGACGCTGCGCG 1

RESULT 7

LOCUS A1113131 315 bp mRNA linear EST 11-FEB-1999

DEFINITION UI-R-C2p-nr-f-06-0-UI.s1 UI-R-C2p Rattus norvegicus cDNA clone

ACCESSION UI-R-C2p-nr-f-06-0-UI 3', mRNA sequence.

VERSION A1113131

KEYWORDS EST.

SOURCE

ORGANISM Norway rat.

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE

1 (bases 1 to 315)

Bonaldo,M.F., Lennon,G. and Soares,M.B.

Normalization and subtraction: two approaches to facilitate gene

JOURNAL

Genome Res. 6 (9), 791-806 (1996)

COMMENT

Contact: Soares, MB

Program for Rat Gene Discovery and Mapping

University of Iowa

451 Eckstein Medical Research Building Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: meares@blue.weeg.uiowa.edu

Oligo-dt track not found, Not 1 site shown in beginning of sequence

is likely internal to the message. cDNA Library Preparation: M.

Factima Bonaldo, Ph.D. Clone distribution: clones will be available

through Research Genetics The following repetitive elements were

found in this cDNA sequence: 97-125, >GC-rich#low_complexity

Seq primer: M13 Forward

Location/Qualifiers

1..315

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/clone="UI-R-C2p-nr-f-06-0-UI"

/clone_1ib="UI-R-C2p"

/dev_stage="adult"

/lab_host="DH10B (Life Technologies)"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker, Site 1: Not I; Site 2: Eco RI. The UI-R-C2p

library is a subtracted library derived from the UI-R-C1

library, which is a subtracted library derived from the

UI-R-C0 library. The UI-R-C0 library consisted of a

mixture of individually tagged normalized libraries

constructed from adult placenta, adult lung, brain, liver,

kidney, heart, spleen, ovary, muscle, 8, 12 and 18-day

embryo. The tag is a string of 3-5 nucleotides present

between the Not I site and the oligo-dt track which allows

identification of the library of origin of a clone within

the mixture. The subtracted library (UI-R-C2p) was

constructed as follows: PCR amplified cDNA inserts from

UI-R-C1 clones from which 3' ESTs had been derived was

used as a driver in a hybridization with the UI-R-C1

library in the form of single-stranded circles. The

remaining single-stranded circles (subtracted library) was

purified by hydroxyapatite column chromatography,

converted to double-stranded circles and electroporated

into DH10B bacteria (Life Technologies) to generate the

UI-R-C2p library. This procedure has been previously

described (Bonaldo, Lennon and Soares, Genome Research 6:

791-806, 1996)"

791-806, 1996"

BASE COUNT 46 a 110 c 116 g 43 t

ORIGIN

Query Match 45.5%; Score 289.4; DB 9; Length 315;

Best Local Similarity 94.9%; Pred. No. 1.5e-52;

Matches 299; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

OY 247 TACACCCGCTTCTGACAGACGAGCCATGCCGAGCGCAAGCCGATCACCAGTTGGTG 306

Db 315 TACACCCGCTTCTGACAGACGAGCCATGCCGAGCGCAAGCCGATCACCAGTTGGTG 256

OY 307 TGTCTCCGCGCAGTGGGCGCCCGCGGCTGCTGCCCAAGCCATCGGCGCGTGAAGTGG 366

Db 255 TGTCTCCGCGCAGTGGGCGCCCGCGGCTGCTGCCCAAGCCATCGGCGCGTGAAGTGG 196

OY 367 TGGCGCCGGAACGAGACCGGATTTCCGCTGCATCCCGATGCTACCGCGCAGCGGGTG 426

Db 195 TGGCGCCGGAACGAGACCGGATTTCCGCTGCATCCCGATGCTACCGCGCAGCGGGTG 136

OY 427 CAGCTGCTGTGCGCGGAGGCGCGCGCGCGCTCGCGCAAGTGCCTGTGGTGGCTCG 486

Db 135 CAGCTGCTGTGCGCGGAGGCGCGCGCGCGCTCGCGCAAGTGCCTGTGGTGGCTCG 76

OY 487 TCGAAGTCAAGCGCTTCACCGCTTCCACCAAGTCGAGCTCAAGAGACTTGGGCGG 546

Db 75 TCGAAGTCAAGCGCTTCACCGCTTCCACCAAGTCGAGCTCAAGAGACTTGGGCGG 16

OY 547 GAGACCGCGCGCGCGG 561

Db 15 GAGACCGCGCGCGCGG 1

RESULT 8

LOCUS BE111224/c 291 bp mRNA linear EST 13-JUN-2000

DEFINITION UI-R-BJ1-auw-b-04-0-UI.s1 UI-R-BJ1 Rattus norvegicus cDNA clone

UI-R-BJ1-auw-b-04-0-UI 3', mRNA sequence.

ACCESSION BE111224

VERSION BE111224.1 GI:8503329

KEYWORDS EST.

SOURCE Norway rat.

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 291)

Bonaldo,M.F., Lennon,G. and Soares,M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)

COMMENT 97044477

Contact: Soares, MB

Program for Rat Gene Discovery and Mapping

University of Iowa

451 Eckstein Medical Research Building Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: meares@blue.weeg.uiowa.edu

Oligo-dt track not found, Not 1 site shown in beginning of sequence

is likely internal to the message. cDNA Library Preparation: M.B.

Soares Lab Clone distribution: clones will be available through

Research Genetics (www.resgen.com) The following repetitive

elements were found in this cDNA sequence: 97-125,

>GC-rich#low_complexity

Seq primer: M13 Forward

POLYA=No.

Location/Qualifiers

1..291

/organism="Rattus norvegicus"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/clone="UI-R-BJ1-auw-b-04-0-UI"

/clone_1ib="UI-R-BJ1"

/lab_host="DH10B (Life Technologies)"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker, Site 1: Not I; Site 2: Eco RI; The UI-R-BJ1

library is a subtracted library derived from the following

tissues: atrium at 16.5 dpc, ventricle at 16.5 dpc, AV

canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc,

AV canal at 15 dpc, ventricle at 13 dpc, and adult heart. For a detailed description of the library from which this clone was derived, please visit our web site at ratser.eng.uiowa.edu. The subtraction has been previously described in Bonaldo, Lemmon and Soares, *Genome Research* 6:791-806, 1996

BASE COUNT	41 a	101 c	109 g	4
ORIGIN				

Query Match	43.0%;	Score 273.4;	DB 10;	Length 291;
Best Local Similarity	96.2%;	Pred. No. 4e-49;		
Matches 280; Conservative	0;	Mismatches 11;	Indels 0;	Gaps 0

Qy	211	CCATCCCCGAGCGCCAAACCGGTCAACGAGTTGGTGTCTCCGCGCAGTGGCGCCGCG	333
Db	291	CGGTCCGCGAGGCCAAACCGGTCAACGAGTTGATGTGTCTCGGCGAAGTGGCGCCGCG	233
Qy	331	CGGCTGCTGCCAACGCCATCGGGGCGGTGAAGTGTGTGGCCCGGAACGACCGGATTTG	393
Db	231	CGGCTGCTGCCAACGCCATCGGGGCGGTGAAGTGTGTGGCCCGGAACGACCGGACTTC	173
Qy	391	CGCTCATCCCGGAGTCGCTACCGCGCGGAGCGGGGTGACGCTGTGGCCCGCGGGGGCGG	453
Db	171	CGCTCATCCCGGAGTCGCTACCGCGCGGAGCGGGGTGACGCTGTGTGCCCGCGGGGGCG	1123
Qy	451	GGCGCGCGCTCGCGCAAGTGTGCTGTGTGGCTCTGTGTCAAGTCAACGCGCTTACCGCG	510
Db	111	GGCGCGCGCTCGCGCAAGTGTGCTGTGTGGCTCTGTGTCAAGTCAACGCGCTTACCGCG	52
Qy	511	TTTCAACAACGATCGGAGACTCAAGACTTTGGGGCGCGAGAACCGCGGGCGG	561
Db	51	TTTCAACAACGATCGGAGACTCAAGACTTTGAACTTGTGAACCGCGGGCGG 1	

RESULT 9	LOCUS	DEFINITION
AQ171546/c	AQ171546	535 bp DNA linear
	HS_3088_B1_H06	MF CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3088 Col=11 Row=P, DNA sequence.

BASE COUNT	98	a	164	c	139	g	132	t	2	others
ORIGIN	/note="Organ: sperm; Vector: pBelobaCl; Bac Clones in E-Coli DH10B"									

Query Match	21.8%	Score 138.8;	DB 17;	length 535;
Best Local Similarity	80.5%;	Pred. No. 4.2e-20;		
Matches 177; Conservative	0;	Mismatches 37;	Indels 6;	Gaps 1;

Db 268 ATGCAGTCCCCACTGAGCCCTGTGTCTGTCTGCTGTGCTGTACACACAGCCCTCCGTGA 209

Db 208 GTGGAGGGCCAGGGGGTGGAGCGTTCAAGAAATGATGCAAGAAATCATCTCCCGAGCTC 149
 Qy 121 GGAAGTACCCCGAGCTCTCTC-----TAGAAACAACAGACCATGAACCGGGCGAG 174
 Db 148 GGAAGTACCCCGAGCTCTCAACGAGCTGAGAAACAAGACCATGAACCGGGCGAG 89
 Qy 175 AATGAGGCGAGCTCCCAACCATCTATGACGCCAAG 214
 Db 88 AACGAGGGGGGCTCTCCACCAACCTTTAGACCAAG 49

RESULT 10	BP523030	254 bp	-mRNA	linear	EST 11-DEC-2000
LOCUS	BP523030				
DEFINITION	UI-R-C3p	-tk-a-03-0-UI.1	UI-R-C3p	Rattus norvegicus	cDNA clone
DEFINITION	UI-R-C3p	-tk-a-03-0-UI.5'			mRNA sequence.

FEATURES
source

```

/ab host="DH10B (Life Technologies)"
/nc=Vector: pYT19-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI. The UI-R-CpD
library is a subtracted library derived from the UI-R-C1
library, which is a subtracted library derived from the
UI-R-C0 library. The UI-R-C0 library consisted of a
mixture of individually tagged normalized libraries
constructed from rat placenta, adult lung, brain, liver,

```


kidney, heart, spleen, ovary, muscle, 8, 12 and 18-day embryo. The tag is a string of 3-5 nucleotides present between the Not I site and the oligo-3T track which allows identification of the library of origin of a clone within the mixture. The subtracted library (UI-R-C2p) was constructed as follows: PCR amplified cDNA inserts from UI-R-C1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with the UI-R-C1 library in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the UI-R-C2p library. This procedure has been previously described (Bonaldi, Lennon and Soares, Genome Research 6: 791-806, 1996).

BASE COUNT 61 a 90 c 53 g 50 t

Query Match 21.4%; Score 136.2; DB 12; Length 254;
Best Local Similarity 79.9%; Pred. No. 1,4e-19;
Matches 175; Conservative 0; Mismatches 38; Indels 6; Gaps 1;

QY 1 ATGACGCGCTCACTAGCGCCGCTGCTATCTGCTAGCGCTGCTGCTGCT 60
DB 35 ATGACGCTCTCACTAGCGCCCTTCTGCTGCTGCTGCTGCTGCTGCT 94
QY 61 GTGAGAGCGCGGCGGTGCGAAGCTTGAAGATGATCCAGAGGTCATCCAGGCTT 120
DB 95 GTGAGAGCGCGGCGGTGCGAAGCTTGAAGATGATCCAGAGGTCATCCAGGCTT 154
QY 121 CGAGAGTACCGCGAGCGCTCTCTCTG-----AGACACACGAGCGCTGAGCGGCGGAG 174
DB 155 ACAGAGTACCGCGAGCGCTCTCTCTGAGAACTATGACACAGCGAGCGCTCTCT 214
QY 175 AATGAGGAGGAGAGCTCCCGACCATCTCCATGAGCGCGAA 213
DB 215 AACGAGGAGGATCCCGCGACCATCTTATGACACCGAA 253

RESULT 11 1612 bp mRNA linear HTC 19-JAN-2002
AK002396 Mus musculus adult male kidney cDNA, RIKEN full-length enriched
LOCUS library: clone:0610009622:homolog to HYPOTHETICAL 15.3 KDA PROTEIN
DEFINITION (FRAGMENT), full insert sequence.

ACCESSION AK002396
VERSION AK002396.1 GI:12832347
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (strain: C57BL/6J) adult male kidney cDNA to mRNA, clone:lib:RIKEN full-length enriched mouse cDNA library
clone:0610009622.

ORGANISM

Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)

JOURNAL MEDLINE 99279253
PUBMED 10349636
2

AUTHORS

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL MEDLINE 20499374
PUBMED 11042159
3

REFERENCE JOURNAL MEDLINE 20499374
PUBMED 11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Kono, H., Akiyama, J., Nishi, K., Katsunari, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,

Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multiplexillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL MEDLINE 20530913
PUBMED 11076861
4

Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Aizawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S.,
Aikawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamataka, I.,
Saito, T., Okazaki, Y., Gotohori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batilov, S., Casavant, T.,
Fleischmann, M., Gaasterland, T., Gissi, C., King, B., Kochwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
Quackenbush, J., Schriml, L.M., Stebbins, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, J.D., Bojunga, N.,
Carninci, P., de Bonaldi, M.F., Brownstein, M.J., Bult, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D.A., Kamliya, M., Lee, N.H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Mombauts, P., Nordone, P.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Sessa, T., Shibata, Y., Storch, K.F., Suzuki, H.,
Toyo-Oka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kotsuki, S.
and Hayashizaki, Y.
Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)

JOURNAL MEDLINE 21085660
PUBMED 11217851
5

(bases 1 to 1612)
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
Arikawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C.,
Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T.,
Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F.,
Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kono, H., Konda, M., Koya, S.,
Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K.,
Numata, R., Ohtsuka, Y., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J.,
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shinkai, T.,
Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,
Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I.,
Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and
Hayashizaki, Y.
Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
Kanagawa 230-0045, Japan (E-mail: genome-research@riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

TITLE
JOURNAL

COMMENT

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for
further details.

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5'-GAGAGGAGAGCGCGCGCACTGCTTTTCTTTTCTTTT 3'], cDNA was
prepared by using trehalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. cDNA went
through one round of normalization to Rot = 5.0. Second strand cDNA
was prepared with the primer adapter of sequence [5'-
GAGAGGAGAGGATCGAGAGCTCAATTAATTAATTAACCCCGCC 3']. cDNA was
cloned with XhoI and SclI. Cloning sites, 5' end: SclI; 3' end:
XhoI. Host: SOLR.

FEATURES
source

Location/Qualifiers
1..1612

AUTHORS	Carninci, P. and Hayashizaki, Y.
TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)
DOI	99279253
MEDLINE	10349636
PMID	10349636
REFERENCE	
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

AUTHORS
Shibata, K., Itoh, M., Aizawa, K., Naeaka, S., Sasaki, N., Carminci, P.,
Komoto, H., Akiyama, Y., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kasaihiwagi, K.,
Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohnita, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunara, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

TITLE
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multiplexillary sequencer

JOURNAL
Genome Res. 10 (11), 1757-1771 (2000)

AUTHORS
Kawai, J., Shingagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Komno, H., Aichi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kasuwa, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kondou, T., Saito, R., Kodera, K., Matuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochia, T., Kuehl, P., Lewis, S., Matsuo, Y., Nikidjo, I., Peesle, G., Quackenbush, J., Schriml, L. M., Staudt, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldwin, I. P., Barsh, G., Blake, J., Botfield, D., Bojunga, N., Cairncell, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Guecinich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Maehima, J., Mazzarelli, J., Mombart, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shihara, S., Strach, K. F., Suzuki, H., Teyo-Oka, K., Wang, K. H., Welz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaku, S.

QY 424 GTGAGCTGCTGTGCCCCCGGGGGCGCGCGCGCTTCCGCAAGTGCCTGTGTGCCC 483
 DB 496 ATCCAGCTGCAAGTGTCAAGAGCGCG--AGCAGCGCACCTTAAATAATCACCCTGTGACG 552
 QY 484 TCGTCAAGTGTCAAGCGCTCACCGCTTCCAGCAACCAAGTGTGAGTCAAGACTTCCGG 543
 DB 553 GCGTCAAGTGTCAAGAGGTACACCGCTCAGCAACAAGTGTCAAGCTTGAAGC 612
 QY 544 CCGAGAACCGCGCGCGCGCGAGAGGTGCAAGCGCGCGCGCGCGCGCGAGCA 601
 DB 613 GTGTGCGCCCGCAAGCGCGCGCGCACCAAGAGAGCGAGAGAGAGCAATCA 670

RESULT 15
 BF607657 669 bp mRNA linear EST 01-APR-2001
 LOCUS BF607657
 DEFINITION M11.000599 Mouse 9-day fetus cDNA library ICRFp522 Mus musculus
 cDNA clone ICRFp522G0852 5', mRNA sequence.
 ACCESSION BF607657
 VERSION BF607657.1 GI:13504149
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 669)
 Yahyawi, M., Hennig, S., Neidhardt, L., Radelof, U., Hermann, B.G.,
 Lehrach, H. and O'Brien, J.
 Detection of a high number of novel genes in a 9-day mouse embryo
 cDNA library normalised by oligonucleotide fingerprinting
 Unpublished (2001)
 JOURNAL Contact: Hennig S
 Laboratory 123, dept. Lehrach
 Max-Planck-Institut fuer Molekulare Genetik
 Insestr.63-73, D-14195 Berlin, Germany
 Tel: +49 30 8413 1612
 Fax: +49 30 8413 1380
 Email: hennig@molgen.mpg.de
 EST's are made from clones being representatives of clone clusters.
 Clone clusters were calculated from oligonucleotide fingerprints.
 PCR Primers
 FORWARD: 5'-GAGCTATTCAGAGTAGTA-3'
 BACKWARD: 5'-TAATAGACTCACTATAGG-3'
 Seq primer: 5'-ATTAGGTGACACTATAG-3'
 High quality sequence stop: 669.
 Location/Qualifiers
 source
 1..669
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="ICRFp522G0852"
 /clone_id="Mouse 9-day fetus cDNA library ICRFp522"
 /tissue_type="whole embryo"
 /dev_stage="embryonic 9-day"
 /lab_host="E.coli, XL1 blue"
 /note="Vector: pSVSPORT1; Site 1: NotI; Site 2: SalI;
 Library preparation by oligo-dT priming of RNA. Clones can
 be ordered from the Resource Center in Berlin,
 http://www.rzpd.de."

BASE COUNT 172 a 202 c 174 g 120 t 1 others
 ORIGIN

Query Match 12.9%; Score 82.2; DB 12; Length 669;
 Best Local Similarity 52.6%; Pred. No. 6.6e-08;
 Matches 283; Conservative 0; Mismatches 234; Indels 21; Gaps 4;

QY 82 GCCTTCAGGATGATGCGACAGAGTCAATCCAGGGCTTGAGAGTACCC--CGAGCCT 138
 DB 128 GCTTTAAATGATGCGACAGAAATCTTTATTCAGATGTGTTAAACCTGTCCGGCA 187
 QY 139 CCTCTGAGAACACAGACCATGAACCGGCGGAGATGAGGACAGACTCC--CAC 195
 DB 188 CACCCAGACAGACAGACCTCTGATCAACAGAGATGAGGACAGCATTTCAATAGC 247

QY 196 CATCCCTATGACCCCAAGATGTGTCCAGTACAGCTGCGCGAGCTGCACTACACCCGC 255
 DB 248 ACTGAGCTGAGATCGAAGACATTCAGTCAAGTGGCTGCAAGGAACTCGGTTCAACAA 307
 QY 256 TTCTGACAGACGCGCCCATGTGCGCAGCGCCCAAGCCGATCACAGTTGTGTGCTCCGC 315
 DB 308 TACATTTGGAAGCGGCAAGTGTGACAGCATCAAGCCTCTTAAAGAGCTGTGTGCGCGGC 367
 QY 316 CAGTGGCGCCCGCGCGCGCTGTGCTCCAAAGCCATCGGCGCG-----TGAG 363
 DB 368 GAGTGTGCTGCCCCCTGCGGCTGTCTCCAACTGSAATCGAGAGAGCTACGAAACAAATAC 427
 QY 364 TGTGTGCGCCCGCAAGAGACCGGATTTCCGCTGCATCCCGGATTCGCTACCGGCGCAGCG 423
 DB 428 TGGAGCCGAGAGAGCTCTCAGAGTGGCGGTGTCAACGACAGACCGCACCCAGAG 487
 QY 424 GTGCACTGTGTGCCCCCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 483
 DB 488 ATCCAGCTGTCAATGTG--AGGACGCGACAGACCGCACCTTAAATAACCGTGTGACG 544
 QY 484 TGTGCAAGTGTCAAGCGCCTCACTGCTTCAACAAACAGTGTGAGCTCAAGACTTCCGG 543
 DB 545 GCGTCAAGTGTCAAGAGGTACACCGCTCAACAAACAGTGTGAGCTCAAGACTTGAAGC 604
 QY 544 CCGAGAACCGCGCGCGCGCGAGAGGTGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCG 601
 DB 605 GTGTGCGCCAGCAAGCGCGCGCGCGCACCAAGAGAGCGGAGAGCCAGCATATCA 662

Search completed: March 29, 2003, 01:28:53
 Job time : 1372.11 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Comphen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 28, 2003, 18:49:00 ; Search time 76.1376 Seconds
(without alignments)
7108.427 Million cell updates/sec

Title: US-09-867-274-3

Perfect score: 636
Sequence: 1 atgcagccctcactagcccc.....agctgagagcgcctactag 636

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 574371 seqs, 425486471 residues

Total number of hits satisfying chosen parameters: 1148742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: Published Applications NA.*
2: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
3: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	636	100.0	636	10	US-09-867-274-3
2	501.6	78.9	759	10	US-09-867-274-1
3	358	56.3	422	10	US-09-864-761-30988
4	232.4	36.5	392	10	US-09-864-761-14440
5	81.6	12.8	954	9	US-10-152-661-430
6	81.6	12.8	954	9	US-09-866-050A-430
7	81.6	12.8	962	9	US-10-152-661-40
8	81.6	12.8	962	9	US-10-152-661-209
9	81.6	12.8	962	9	US-09-866-050A-40
10	81.6	12.8	962	9	US-09-866-050A-209
11	67.2	10.6	900	10	US-09-853-625B-1
12	65.6	10.3	1692	10	US-09-745-286-92
13	44.2	6.9	373	9	US-09-803-719-2375
14	44.2	6.9	426	9	US-09-803-719-2375
15	44.2	6.9	1951	10	US-09-822-830A-91
16	44.2	6.9	3930	10	US-09-847-874A-2
17	44.2	6.9	43058	10	US-09-954-456-292
18	44.2	6.9	43058	10	US-09-954-456-529
19	44.2	6.9	43058	10	US-09-880-107-3950

C	20	44	6.9	174424	10	US-09-867-768A-314	Sequence 314, App
C	21	43	6.8	4020	9	US-09-796-679-5	Sequence 5, Appli
C	22	42.8	6.7	11220	9	US-09-860-846-32	Sequence 32, Appli
C	23	42.8	6.7	11220	10	US-09-861-289-32	Sequence 32, Appli
C	24	42.8	6.7	36778	9	US-09-860-846-5	Sequence 5, Appli
C	25	42.8	6.7	36778	10	US-09-861-289-5	Sequence 5, Appli
C	26	42.6	6.7	380	9	US-09-803-719-2374	Sequence 2374, Ap
C	27	42.6	6.7	8833	10	US-09-880-107-3272	Sequence 3272, Ap
C	28	42.4	6.7	5828	9	US-09-430-029-1	Sequence 1, Appli
C	29	42.2	6.6	4039	10	US-09-205-448-7	Sequence 7, Appli
C	30	41.6	6.5	375	9	US-09-803-719-2373	Sequence 2373, Ap
C	31	41.6	6.5	1230	9	US-09-894-844-114	Sequence 114, App
C	32	41.4	6.5	953	10	US-09-452-239-45	Sequence 45, Appli
C	33	41.4	6.5	3537	9	US-09-858-546-1	Sequence 1, Appli
C	34	40.8	6.4	495	10	US-09-863-824-3	Sequence 3, Appli
C	35	40.8	6.4	1395	10	US-09-863-824-1	Sequence 1, Appli
C	36	40.8	6.4	2898	10	US-09-737-149-3	Sequence 3, Appli
C	37	40.8	6.4	2916	10	US-09-737-149-5	Sequence 5, Appli
C	38	40.8	6.4	3132	10	US-09-737-149-7	Sequence 7, Appli
C	39	40.6	6.4	3713	9	US-09-927-827-19	Sequence 19, Appli
C	40	40.6	6.4	4098	10	US-09-962-436-37	Sequence 37, Appli
C	41	40.6	6.4	4368	9	US-09-981-353-15	Sequence 15, Appli
C	42	40.6	6.4	5427	9	US-09-981-353-159	Sequence 159, App
C	43	40.4	6.4	33023	10	US-09-880-107-3350	Sequence 3350, Ap
C	44	40.2	6.3	5025	10	US-09-960-253-176	Sequence 176, App
C	45	40	6.3	2658	10	US-09-815-242-4035	Sequence 4035, Ap

ALIGNMENTS

RESULT 1									
US-09-867-274-3									
Sequence 3, Application US/09867274									
Patent No. US20020106650A1									
GENERAL INFORMATION:									
APPLICANT: Pastely, Christopher									
APPLICANT: Gao, Yongming									
TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof									
FILE REFERENCE: 01017/37428									
CURRENT FILING DATE: 2001-05-29									
PRIOR APPLICATION NUMBER: US 60/208,550									
PRIOR FILING DATE: 2000-06-01									
PRIOR APPLICATION NUMBER: US 60/223,542									
PRIOR FILING DATE: 2000-08-04									
NUMBER OF SEQ ID NOS: 25									
SOFTWARE: PatentIn version 3.0									
SEQ ID NO 3									
LENGTH: 636									
TYPE: DNA									
ORGANISM: Mus musculus									
US-09-867-274-3									
Query Match									
Best Local Similarity 100.0%; Score 636; DB 10; Length 636;									
Matches 636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	ATGCAGCCCTCAGTACGCCCCGCTGCTCATCTGCTCACTTGTGACGCTGCTTGTGCT	60						
DB	1	ATGCAGCCCTCAGTACGCCCCGCTGCTCATCTGCTCACTTGTGACGCTGCTTGTGCT	60						
QY	61	GTGAGGGGCGAGGGGTGGCAAGCTTTCAGATGATGACAGAGGTCATCCCGAGGCTT	120						
DB	61	GTGAGGGGCGAGGGGTGGCAAGCTTTCAGATGATGACAGAGGTCATCCCGAGGCTT	120						
QY	121	GGAGAGTACCCGAGCCTCTCTTGAAGAACACGACCATGAAACCGGCGGAGAAATGA	180						
DB	121	GGAGAGTACCCGAGCCTCTCTTGAAGAACACGACCATGAAACCGGCGGAGAAATGA	180						
QY	181	GGCAGACCTCCCGACCATCTTATGACGCAAAAGTGTCTCGAGTACAGCTGCCGAG	240						
DB	181	GGCAGACCTCCCGACCATCTTATGACGCAAAAGTGTCTCGAGTACAGCTGCCGAG	240						

QY 241 CTGACTACACCGCTTCTTGACAGACGCGCCATGCGCGAGCGCCGAAGCCGGTCAACCGAG 300
DB 241 CTGACTACACCGCTTCTTGACAGACGCGCCATGCGCGAGCGCCGAAGCCGGTCAACCGAG 300
QY 301 TTGGTGTGCTCCGGCCAGTGGCGCGCGCGCGCTGCTGCCCAACGCGCATCGAGCGCGTG 360
DB 301 TTGGTGTGCTCCGGCCAGTGGCGCGCGCGCGCTGCTGCCCAACGCGCATCGAGCGCGTG 360
QY 361 AAGTGTGGCGCGCGAGACCGGATTTCCGTGCAATCCCGATCGCTACCGCGCGAG 420
DB 361 AAGTGTGGCGCGCGAGACCGGATTTCCGTGCAATCCCGATCGCTACCGCGCGAG 420
QY 421 CGGGTGCAGCTGTGTGCTCCGGGGGGCGCGCGCGCGCTGCGCGCAAGTGGCTGTGTG 480
DB 421 CGGGTGCAGCTGTGTGCTCCGGGGGGCGCGCGCGCGCTGCGCGCAAGTGGCTGTGTG 480
QY 481 GCCTGTGCAAGTGCAGACCGCTTCAACCAACAGTGGAGCTCAAGAGCTTC 540
DB 481 GCCTGTGCAAGTGCAGACCGCTTCAACCAACAGTGGAGCTCAAGAGCTTC 540
QY 541 GGGCGGAGACCGCGCGCGCGCGAGAGGTGCAAGCGCGCGCGCGCGCGCGAGAGCC 600
DB 541 GGGCGGAGACCGCGCGCGCGCGAGAGGTGCAAGCGCGCGCGCGCGCGCGAGAGCC 600
QY 601 AAGCCAAACGAGCGGAGCTGGAAGCGCTACTAG 636
DB 601 AAGCCAAACGAGCGGAGCTGGAAGCGCTACTAG 636

RESULT 2

US-09-867-274-1
Sequence 1, Application US/09867274
Patent No. US20020106850A1
GENERAL INFORMATION:
APPLICANT: Paszty, Christopher
APPLICANT: Gao, Yonming
TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
FILE REFERENCE: 01017/37428
CURRENT APPLICATION NUMBER: US/09/867,274
PRIOR FILING DATE: 2001-05-29
PRIOR APPLICATION NUMBER: US 60/208,550
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: US 60/223,542
PRIOR FILING DATE: 2000-08-04
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 759
TYPE: DNA
ORGANISM: Homo sapiens
US-09-867-274-1

Query Match 78.9%; Score 501.6; DB 10; Length 759;
Best Local Similarity 87.5%; Pred. No. 4,4e-126;

Matches 562; Conservative 0; Mismatches 74; Indels 6; Gaps 1;

QY 1 ATGCAAGCCCTCACTAGCCCGCTGCTATCTGCTCTTGTGACGCTGCTTGTGCT 60
DB 37 ATGCAAGCTCCCACTAGCGCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 96
QY 61 GTGAGGGGCGAGGGGTGGCAAGCTTCAAGATATGCAAGAGGTGATCCCGAGGCTT 120
DB 97 GTGAGGGGCGAGGGGTGGCAAGCTTCAAGATATGCAAGAGGTGATCCCGAGGCTT 156
QY 121 GGAAGTATCCCGAGGCTCTTC-----TGAGAACAAACGAGCATGAACCGGGCGAG 174
DB 157 GGAAGTATCCCGAGGCTCTTC-----TGAGAACAAACGAGCATGAACCGGGCGAG 216
QY 175 AATGAGGCGAGACCTCCCGACCATCTTATAGCGCAAGAGTGTGTCGAGTACAGCTGC 234
DB 217 AATGAGGCGAGACCTCCCGACCATCTTATAGCGCAAGAGTGTGTCGAGTACAGCTGC 276

QY 235 CGGAGCTGCACTACACCGCTTCTTGACAGACGCGCCATGCGCGAGCGCCGAAGCCGGTGC 294
DB 277 CGGAGCTGCACTTACACCGCTTCTTGACAGACGCGCCATGCGCGAGCGCCGAAGCCGGTGC 336
QY 295 ACCGAGTGTGTGCTCCGGCCAGTGGCGCGCGCGCGCTGCTGCCCAACGCGCATCGAG 354
DB 337 ACCGAGTGTGTGCTCCGGCCAGTGGCGCGCGCGCGCTGCTGCCCAACGCGCATCGAG 396
QY 355 CGCGTGAAGTGTGTGCGCGCGCGCGAGCGGATTTCCGTGCAATCCCGATCGCTACCGC 414
DB 397 CGCGTGAAGTGTGTGCGCGCGCGAGCTAGTGGGCCCACTTCCGTGCAATCCCGATCGCTACCGC 456
QY 415 GCGAGCGGAGTCACTGCTGTGCTCCCGGGGGCGCGCGCGCTGCGCGCAAGTGGT 474
DB 457 GCGAGCGGAGTCACTGCTGTGCTCCCGGGGGCGCGCGCGCTGCGCGCAAGTGGT 516
QY 475 TTGTGTGCTGTGTGCAAGTGCAGAGGCTTCAACCAACAGTGGAGTCTAG 534
DB 517 CTGTGTGCTGTGTGCAAGTGCAGAGGCTTCAACCAACAGTGGAGTCTAG 576
QY 535 GACTTGGGCGCGGAGACCGCGCGCGCGCGAGAGGTGCGCAAGCGCGCGCGCGCGCGCG 594
DB 577 GACTTGGGCGCGGAGACCGCGCGCGCGAGAGGTGCGCAAGCGCGCGCGCGCGCGCG 636
QY 595 GGAGCCAAACGAGCGGAGCTGGAAGCGCTACTAG 636
DB 637 GGAGCCAAACGAGCGGAGCTGGAAGCGCTACTAG 678

RESULT 3

US-09-864-761-30988/c
Sequence 30988, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wenheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecm1ca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
PRIOR FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 30988
LENGTH: 422
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC003098.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.62
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.6
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.49
OTHER INFORMATION: EST HUMAN HIT: BE613498.1, EVALU 5.90e-02
OTHER INFORMATION: SWISSPROT HIT: P45646, EVALU 4.90e-01
OTHER INFORMATION: NT HIT: AF074705.1, EVALU 1.00e+00
US-09-864-761-30988

Query Match 56.3%; Score 358; DB 10; Length 422;
Best Local Similarity 90.5%; Pred. No. 1.8e-87;
Matches 382; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 215 ATGTGTCAGTACAGTCCGAGAGTGCATACACCGCTTCTGACAGAGCCCAT 274
DB 422 ACCTGTCAGTACAGTCCGAGAGTGCATACACCGCTTCTGACAGAGCCCAT 363
QY 275 GCGCAGCGCCAGACCGGTCACCGAGTGTGCTCCGCGCAGTGGCGCGCGC 334
DB 362 GCGCAGCGCCAGACCGGTCACCGAGTGTGCTCCGCGCAGTGGCGCGCGC 303
QY 335 TGCTGCCCAAGCGCATGGCGGGGCTGAAGTGTGGCGCGCGCAGACCGATTCCGCT 394
DB 302 TGCTGCCCAAGCGCATGGCGGGGCTGAAGTGTGGCGCGCGCAGACCGATTCCGCT 243
QY 395 GCATCCCGGATTCGCTACCGCGCGCAGCGGCTGAGCTGTGCTCCCGGGGCGCGCGC 454
DB 242 GCATCCCGGATTCGCTACCGCGCGCAGCGGCTGAGCTGTGCTCCCGGGGCGCGCGC 183
QY 455 CGCGCTCGCGCAGAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 514
DB 182 CGCGCTCGCGCAGAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 123
QY 515 ACAACGAGTGGAGCTCAAGAGATTGGGCGCGAGACCGCGCGCGCAGAGAGGCTCGCA 574
DB 122 ACAACGAGTGGAGCTCAAGAGATTGGGCGCGAGACCGCGCGCGCAGAGAGGCTCGCA 63
QY 575 AGCG 634
DB 62 AGCG 3
QY 635 AG 636
DB 2 AG 1

RESULT 4
US-09-864-761-14440/c
Sequence 14440, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aemica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 14440
LENGTH: 392
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC003098.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.62
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.6
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.49
US-09-864-761-14440

Query Match 36.5%; Score 232.4; DB 10; Length 392;
Best Local Similarity 90.5%; Pred. No. 1.1e-53;
Matches 248; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 363 GTGTGGGCG 422
DB 392 GTGTGGGCG 333
QY 423 GGTGAGCTGTGTGCTCCCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 482
DB 332 GGTGAGCTGTGTGCTCCCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 273
QY 483 CTGTGCAAGTGCAGCGCTTCACCGCTTCACAGCAGTGCAGAGCTCAAGAGCTTCGG 542
DB 272 CTGTGCAAGTGCAGCGCTTCACCGCTTCACAGCAGTGCAGAGCTTCGG 213
QY 543 GCGCGAGCG 602
DB 212 GACCGAGCG 153
QY 603 AGCGAAGCGCGAGTGCAGAGAGCGCTTACG 636
DB 152 AGCGAAGCGCGAGTGCAGAGAGCGCTTACG 119

RESULT 5

US-10-152-661-430
 ; Sequence 430, Application US/10152661
 ; Publication No. US2003022835A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Watson, James D.
 ; APPLICANT: Strachan, Lorna
 ; APPLICANT: Sleeman, Matthew
 ; APPLICANT: Onrust, Rene
 ; APPLICANT: Murison, James G.
 ; APPLICANT: Kumble, Krishanand D.
 ; TITLE OF INVENTION: Compositions Isolated From Skin Cells
 ; TITLE OF INVENTION: and Methods for Their Use
 ; FILE REFERENCE: 11000.1011C5
 ; CURRENT APPLICATION NUMBER: US/10/152,661
 ; PRIOR FILING DATE: 2002-05-20
 ; PRIOR APPLICATION NUMBER: 09/866,050
 ; PRIOR FILING DATE: 2001-05-24
 ; PRIOR APPLICATION NUMBER: 60/221,232
 ; PRIOR FILING DATE: 2000-07-25
 ; PRIOR APPLICATION NUMBER: 60/206,650
 ; PRIOR FILING DATE: 2000-05-24
 ; PRIOR APPLICATION NUMBER: 09/312,283
 ; PRIOR FILING DATE: 1999-05-14
 ; PRIOR APPLICATION NUMBER: PCT/NZ99/00051
 ; PRIOR FILING DATE: 1999-04-29
 ; PRIOR APPLICATION NUMBER: 09/188,930
 ; PRIOR FILING DATE: 1998-11-09
 ; PRIOR APPLICATION NUMBER: 09/069,726
 ; PRIOR FILING DATE: 1998-04-29
 ; NUMBER OF SEQ ID NOS: 725
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 430
 ; LENGTH: 954
 ; TYPE: DNA
 ; ORGANISM: Mouse
 US-10-152-661-430

Query Match 12.8%; Score 81.6; DB 9; Length 954;
 Best Local Similarity 52.6%; Pred. No. 4e-13;

Matches 283; Conservative 0; Mismatches 234; Indels 21; Gaps 4;

QY 82 GCCTTCAGAGTATGATCCACAGAGTATCCAGAGGCTTGAGAGTACCC---CGAGCCT 138
 DB 105 GCTTTAAATAATGATGCGACAGAAATCTTTATTCACATGTGTTAAACCTGTCTCCGCA 164
 QY 139 CCTCTGAGAACAAACAGACCATGAAACCGGCGGAGATGAGGACAGACTCC---CAC 195
 DB 165 CACCCACAGACAGACAGACCCCTGATCAAGCCAGAAATGAGGACAGGCAATTCAGTAGC 224
 QY 196 CATCCCTTATGACCCAAAGATGTGTCGAGTACAGCTGCCGAGCTGCACTAACCCGC 255
 DB 225 ACTGACTGATCGAATAAGTCGAGTTCAAGTGGCTGCGAGAACTGCGCTCCACCAA 284
 QY 256 TTCTCGAAGAGCGGCCATGCGGAGCGCCAGGCGGCTGAGTGTGTCTCCGGC 315
 DB 285 TACATTTGAGACGGCGAGTGCACAGCAATCAGCCCTTGAAGGAGCTGTGTGCGGGC 344
 QY 316 CAGTGCAGCCCGCGCGCTGCTGCCAAGCCCATTCGGGCGC-----TGAAG 363
 DB 345 GAGTGTCTTGGCCCTGCGGTGCTTCCCACTGATCGAGAGAGGCTACGAAACAAATAC 404
 QY 364 TGTGTGCGCCCGAAGCGAGATTCGCTGTCATCCCGATTCCTCAACCGCGCAGCGG 423
 DB 405 TGGAGCCGAGAGAGCTTCAGAGTGTGCGGTGTGTCACAAAGACCGCAGCACAAGG 464
 QY 424 GTGAGCTGTGTGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 483
 DB 465 ATCCAGCTGAGTGTGAGAGAGCGC---AGCAGCGCGCACTTCAAAATACCGGTGTACG 521
 QY 484 TGTGCAAGTGAAGCGCTCAGCCGCTTCCAAACCAAGTGTGAGAGCTTCAAGAGCTTGGG 543

DB 522 GCGTGAAGTGAAGAGGTACACCCGTGAGCAACAGATGCTCAGCAACTTTGAAGC 581
 QY 544 CCGAGAGCCCG 601
 DB 582 GTGTGCGCCAGCCAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 639

RESULT 6

US-09-866-050A-430
 ; Sequence 430, Application US/09866050A
 ; Publication No. US2003004071A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Watson, James D.
 ; APPLICANT: Strachan, Lorna
 ; APPLICANT: Sleeman, Matthew
 ; APPLICANT: Onrust, Rene
 ; APPLICANT: Murison, James G.
 ; APPLICANT: Kumble, Krishanand D.
 ; TITLE OF INVENTION: Compositions Isolated From Skin Cells
 ; TITLE OF INVENTION: and Methods for Their Use
 ; FILE REFERENCE: 11000.1011C4U
 ; CURRENT APPLICATION NUMBER: US/09/866,050A
 ; PRIOR FILING DATE: 2001-05-24
 ; NUMBER OF SEQ ID NOS: 725
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 430
 ; LENGTH: 954
 ; TYPE: DNA
 ; ORGANISM: Mouse
 US-09-866-050A-430

Query Match 12.8%; Score 81.6; DB 9; Length 954;
 Best Local Similarity 52.6%; Pred. No. 4e-13;

Matches 283; Conservative 0; Mismatches 234; Indels 21; Gaps 4;

QY 82 GCCTTCAGAGTATGATCCACAGAGTATCCAGAGGCTTGAGAGTACCC---CGAGCCT 138
 DB 105 GCTTTAAATAATGATGCGACAGAAATCTTTATTCACATGTGTTAAACCTGTCTCCGCA 164
 QY 139 CCTCTGAGAACAAACAGACCATGAAACCGGCGGAGATGAGGACAGACTCC---CAC 195
 DB 165 CACCCACAGACAGACAGACCCCTGATCAAGCCAGAAATGAGGACAGGCAATTCAGTAGC 224
 QY 196 CATCCCTTATGACCCAAAGATGTGTCGAGTACAGCTGCCGAGCTGCACTAACCCGC 255
 DB 225 ACTGACTGATCGAATAAGTCGAGTTCAAGTGGCTGCGAGAACTGCGCTCCACCAA 284
 QY 256 TTCTCGAAGAGCGGCCATGCGGAGCGCCAGGCGGCTGAGTGTGTCTCCGGC 315
 DB 285 TACATTTGAGACGGCGAGTGCACAGCAATCAGCCCTTGAAGGAGCTGTGTGCGGGC 344
 QY 316 CAGTGCAGCCCGCGCGCTGCTGCCAAGCCCATTCGGGCGC-----TGAAG 363
 DB 345 GAGTGTCTTGGCCCTGCGGTGCTTCCCACTGATCGAGAGAGGCTACGAAACAAATAC 404
 QY 364 TGTGTGCGCCCGAAGCGAGATTCGCTGTCATCCCGATTCCTCAACCGCGCAGCGG 423
 DB 405 TGGAGCCGAGAGAGCTTCAGAGTGTGCGGTGTGTCACAAAGACCGCAGCACAAGG 464
 QY 424 GTGAGCTGTGTGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 483
 DB 465 ATCCAGCTGAGTGTGAGAGAGCGC---AGCAGCGCGCACTTCAAAATACCGGTGTACG 521
 QY 484 TGTGCAAGTGAAGCGCTCAGCCGCTTCCAAACCAAGTGTGAGAGCTTCAAGAGCTTGGG 543
 DB 522 GCGTGAAGTGAAGAGGTACACCCGTGAGCAACAGATCGAGCAACTTTGAAGC 581
 QY 544 CCGAGAGCCCG 601
 DB 582 GTTGTGCGCCAGCCAGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 639

RESULT 7

```
US-10-152-661-40
; Sequence 40, Application US/10152661
; Publication No. US20030022835A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c5
; CURRENT APPLICATION NUMBER: US/10/152,661
; PRIOR FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: 09/866,050
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/221,232
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: 60/206,650
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: 09/312,283
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: PCT/NZ99/00051
; PRIOR FILING DATE: 1999-04-29
; PRIOR APPLICATION NUMBER: 09/188,930
; PRIOR FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: 09/069,726
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 962
; TYPE: DNA
; ORGANISM: Mouse
US-10-152-661-40

Query Match      12.8%; Score 81.6; DB 9; Length 962;
Best Local Similarity 52.6%; Pred. No. 4e-13;
Matches 283; Conservative 0; Mismatches 234; Indels 21; Gaps 4;

QY 82 GCCTTGAAGATGATGCGACAGAGGTGATCCAGGGCTTGAGAGTACCC--CGAGCCT 138
DB 113 GCTTTAAATAATGATGCGACAGAAATCCTTATTCACATGTGTTAAACCTGTCGGGCA 172
QY 139 CCTCTGAGAACACACAGACATGAAACCGGGCGGAAATGAGAGCAGACCTCCC--CAC 195
DB 173 CACCCAGAGCAACAGACACCTGATCAAGACAGAGATGAGAGGAGGCAATTTCACTAGC 232
QY 196 CATCCCTATGACCCCAAGATGTGTCCGAGTACAGCTGCCGAGCTGCACTACACCCGC 255
DB 233 ACTGAGACTGATGAGAAACATGATGAGTTCAAGTGGCTGCGAGGAACTGCGGTCAACAA 292
QY 256 TTCTTGAACAGAGCCCATGTCGCGAGGCGCAAGCCGGTCAACGAGTTGTGTCTCCGGC 315
DB 293 TACATTTGCGAGCGGCAAGTGCACACGATCAGCCCTGTGAAGAGCTGTGTGCGGGGC 352
QY 316 CAGTGGGCGCCCGCGGGCTGTGCGCAAGCGATCGGGGCG-----TGAAG 363
DB 353 GAGTCTTGCCCTGCGCGGTCTTCCAACTGATGAGAGAGGCTACCGAACAAAGTAC 412
QY 364 TGTGTGCGCCGAGACCGGATTTCCGCTGATCCCGGATCGTACCGCGCGCAGCGG 423
DB 413 TGGAGCGGAGAGGCTGTCAAGAGTGGCGGTGTGTCAACGACAAAGCGCCACCGAGG 472
QY 424 GTGCAAGCTGTGTGCGCGGGGGCGCGCGCGCTCGCGCAAGGTGCGTGTGTGGCC 483
DB 473 ATCAGGTGAGGTGTCAAGAGCGC---AGCAGCGGCACTTACAAATCAACCTGTGTACG 529
QY 484 TGTGCAAGTGAAGGCGCTCACCGGCTTCAACACAGTCCGAGAGCTCAAGAGCTTCGGG 543
DB 530 GCTTGAAGTGAAGAGGTACACCTCGTCAAGACAAAGTCCAGCCACCACTTTGAAG 589
```

```
QY 544 CCGAGACCGCGCGCCGACAGAGGTGCGAAGCGCGCGCCCGCGCGGAGGACCA 601
DB 590 GTGTGCGACGACGACCGCGCCGACGACCAAGAGAGCGGAAAGAGCCAGCAATCCA 647

RESULT 8
US-10-152-661-209
; Sequence 209, Application US/10152661
; Publication No. US20030022835A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c5
; CURRENT APPLICATION NUMBER: US/10/152,661
; PRIOR FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: 09/866,050
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/221,232
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: 60/206,650
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: 09/312,283
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: PCT/NZ99/00051
; PRIOR FILING DATE: 1999-04-29
; PRIOR APPLICATION NUMBER: 09/188,930
; PRIOR FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: 09/069,726
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 209
; LENGTH: 962
; TYPE: DNA
; ORGANISM: Mouse
US-10-152-661-209

Query Match      12.8%; Score 81.6; DB 9; Length 962;
Best Local Similarity 52.6%; Pred. No. 4e-13;
Matches 283; Conservative 0; Mismatches 234; Indels 21; Gaps 4;

QY 82 GCCTTGAAGATGATGCGACAGAGGTGATCCAGGGCTTGAGAGTACCC--CGAGCCT 138
DB 113 GCTTTAAATAATGATGCGACAGAAATCCTTATTCACATGTGTTAAACCTGTCGGGCA 172
QY 139 CCTCTGAGAACACACAGACATGAAACCGGGCGGAAATGAGAGCAGACCTCCC--CAC 195
DB 173 CACCCAGAGCAACAGACACCTGATCAAGACAGAGATGAGAGGAGGCAATTTCACTAGC 232
QY 196 CATCCCTATGACCCCAAGATGTGTCCGAGTACAGCTGCCGAGCTGCACTACACCCGC 255
DB 233 ACTGAGACTGATGAGAAACATGATGAGTTCAAGTGGCTGCGAGGAACTGCGGTCAACAA 292
QY 256 TTCTTGAACAGAGCCCATGTCGCGAGGCGCAAGCCGGTCAACGAGTTGTGTCTCCGGC 315
DB 293 TACATTTGCGAGCGGCAAGTGCACACGATCAGCCCTGTGAAGAGCTGTGTGCGGGGC 352
QY 316 CAGTGGGCGCCCGCGGGCTGTGCGCAAGCGATCGGGGCG-----TGAAG 363
DB 353 GAGTCTTGCCCTGCGCGGTCTTCCAACTGATGAGAGAGGCTACCGAACAAAGTAC 412
QY 364 TGTGTGCGCCGAGACCGGATTTCCGCTGATCCCGGATCGTACCGCGCGCAGCGG 423
DB 413 TGGAGCGGAGAGGCTGTCAAGAGTGGCGGTGTGTCAACGACAAAGCGCCACCGAGG 472
QY 424 GTGCAAGCTGTGTGCGCGGGGGCGCGCGCGCTCGCGCAAGGTGCGTGTGTGGCC 483
```

Db	473	ATCCAGCTGCAGTGTCAAGAGGCG--AGCAAGCGCACTTCAAAATATCACTGTTGTCAG	529
Qy	484	TTGTGCAAGTACAAAGCGCTCACCCTGTTCCAAACAAGTGGAGTCAAGACTTGGG	533
Db	530	GCTGTCAAGTGTCAAGAGGTATACCTGTCAGACAAACAGTCTACGACCAACTTTGAAAGC	589
Qy	544	CCGAGAACCGCGCGCGCCACAGAGGTGTGCAAGCCGCGCGCGCGCGCGGAGGCTA	601
Db	590	GTGTGCGCAGCCAAAGCCCGCCGATCCACAGAGGCGGAAGAGGACGAGCAATTC	647

RESULT 9
US-09-86

```

? Sequence 40, Application US/098666050A
? Publication No. US20030040471A1
? GENERAL INFORMATION:
? APPLICANT: Watson, James D.
? APPLICANT: Strachan, Lorna
? APPLICANT: Sleeman, Matthew
? APPLICANT: Onrust, Rene
? APPLICANT: Murison, James G.
? APPLICANT: Kumble, Krishanand D.
? TITLE OF INVENTION: Compositions Isolated From Skin Cells
? TITLE OF INVENTION: and Methods for Their Use
? FILE REFERENCE: 11000.1011c4U
? CURRENT APPLICATION NUMBER: US/09/866,050A
? CURRENT FILING DATE: 2001-05-24
? NUMBER OF SEQ ID NOS: 725
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 40
? LENGTH: 9662
? TYPE: DNA
? ORGANISM: Mouse
? US-09-866-050A-40

```

Query Match	Score	DB	Length
12.8%	81.6	9	962

Best local similarity 52.08; Freq. NO. 46-13;
Matches 283; Conservative 0; Mismatches 234; Indels 21; Gaps 4

QY	82	GCTTCACAGATGATGTCACAGAGGTATCCCAAGGCTGGAGAGTACCC---CGAGCCT	138
Db	113	GCTTTTAAATATGATGTCACAGAAATCCTTTATTACATGTTGTTAACTGTCCGGCA	172
QY	139	CCTCTCGAACAACACAGACCATTAACCGGGCGAGAAATGAGGACAGCTCCC--CAC	195
Db	173	CACCCACAGACAACACAGACCTTAATCAAGCCACGAATGAGGCAAGCATTTTCAGTAGC	232
QY	196	CATCCCTATGACGGCAAAAGATGTGTCCGAGTACAGCTGCCGAGGCTGCACACTACACCCG	255
Db	233	ACTGACCTGGATTCGAAACAGTGCAGTTTCACTAGTGGCTGACAGGGAACTGGCGGTCCACAAA	292
QY	256	TTTCCTGACAGACGGCCCTATGCGGAGCGCCAAAGCCGGTCAACGAGTTGTGTCTCCGC	315
Db	293	TACATTTTCGACCGGCAGTAGCACAGATCAAGCCTCTGAAGAGAGCTGTGTGTCCGGCG	352
QY	316	CAGTGCGGCCCCGGCGGGCTGCTGCCAAGCCCATCGGGCGG-----TGAG	363
Db	353	GAGTGTTCGCCCTGCGCGGTCTTCCAACTGATTCGAGAGAGGCTACCGAACAAGTAC	412
QY	364	TGTGTGGGCCCGGAACGGACCGGATTTTCCGCTGCATTCGCCGATGCTTACCGCGCGACGGG	423
Db	413	TGAGACCGAGAGACTCTCAGAAATGTGGGTGTCTCAACGACAAGACGCGCACCCAGAGG	472
QY	424	GTGACAGTGTGTGCCCCGGGGGGCGCGGCGCGGCTTCGCGCAAGTGTGCTTGGTGGCC	483
Db	473	ATCCAGCTGCAGTGTCAAGAGAGGC---AGCAGCGGCACCTTACAAATATCACCGTGGTACG	529
QY	484	TCTGTCAAATGCAAGCGGCTTCAACCGCTTCCACAACCATGTGAGTCAAAGGACTTTCGGG	543
Db	530	GGGTCAAAGTCAAGAGGTACACCGGTACGACCAAGAGTCAAGCCCAATCTTGAAGC	589
QY	544	CCGAGACCGGCGCGCGCAAGAGGTTCGAAGCCCGCGCCCGCGCCCGGAGGCCA	601

Db 590 GTGTCCGACGCCAAGCCCGCCAGCACCAAGAGAGCGGAAGAGAGCCAGCAATCCA 647

RESULT 10

```

US-09-866-050A-209
/ Sequence 209, Application US/09866050A
/ Publication No. US20030040471A1
/ GENERAL INFORMATION:
/ APPLICANT: Watson, James D.
/ APPLICANT: Strachan, Lorna
/ APPLICANT: Sleeman, Matthew
/ APPLICANT: Ornsted, Rene
/ APPLICANT: Morrison, James G.
/ APPLICANT: Kumble, Krishanand D.
/ TITLE OF INVENTION: Compositions Isolated From Skin Cells
/ TITLE OF INVENTION: and Methods for Their Use
/ FILE REFERENCE: 11000.1011c4U
/ CURRENT APPLICATION NUMBER: US/09/866,050A
/ CURRENT FILING DATE: 2001-05-24
/ NUMBER OF SEQ ID NOS: 725
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 209
/ LENGTH: 562
/ TYPE: DNA
/ ORGANISM: Mouse
/ US-09-866-050A-209

```

Query Match 12.8%; Score 81.6; DB 9; Length 962;

Best Local Similarity 52.6%; Pred. No. 4e-13;
Matches 283; Conservative 0; Mismatches 234; Indels 21; Gaps 4

D	113	GCCTTTAAAAATGATGCCACGAAATCTTTATTTCACATGTGGTTAAACCTGTGCCGCA	172
Q	82	GCCTTCAGGAATGATGCCACGAGAGTCATCCAGGGCTTGGAGAGTACC---	CGAGCT 138
Q	139	CTCTCTGAGAAACACCAAGACCATTAACCCGGGCGGAAATATGAGGGTACCTCC---	CAC 159
D	173	CACCCACGACGACACAGCACCTCTAAATCAAGCCAGAAATGAGGGAGGCAATTTTCAGTAGC	232
Q	196	CATCCCTATGACGACCAAGATGTCCGAGTACAGCTGCACGAGACTCATTACACCGC	255
D	233	ACTGACATGGATCGAAACAGTCCAGTTCAAGTGGGTGACGAGGAACCTGGGTCCACCAA	292
Q	256	TTCTTCAGACAGACGCGCCATGCGCGACGCGCCAGCCGGTCAACGAGTTGGTGCTCCGGC	315
D	293	TACATTTCCGAGGCGCACTGTGACACAGATCAGCCCTCTAAGAGACTGTGTGCGCGGCG	352
Q	316	CAGTGCAGCCCCGCGCGCTGTGCCCAACGACCATCGGGCGCG-----TGAAG	363
D	353	GAGTGTGTCCTCCCTCCGAGTCTTCCCACTGGATGAGAGAGGCTACGAAACAAATGAC	412
Q	364	TGGTGGCGCCCGAACGACCGGATTTCCGCTGCATCCCGGATGCGTACCGCGCGACGCG	423
D	413	TGGAAGCCGAGAGACTCTCAGAGATGCGGTGTGTCAACGACAAACCGCACCCACAGG	472
Q	424	GTCACACTGTGTGCCCCGGGGGGCGGCGCGCTGCACGCAAGTGTGCTGTGGGC	483
D	473	ATCCAACTTCAAGTGTCAAGAGCGC--AGACCGGCACTTAATAATACCGTGGTACG	522
Q	484	TCTGCAATGTCAAGCGCTTACACCCGCTTCCACACCACTGTGAGCTCAAGACCTTGCGG	543
D	530	GCATGCAATGTCAAGAGGTACACCCGTCAACACAAAGATTCAGCCACAACTTTGAAGC	589
Q	544	CCGGAATCCGCGCGGCGCAAAAGGTGTGCAAGCCCGCGCGCGCGCGCGCGCGCGCA	601
D	590	GTGTGCCACGACCAACCCCGCCACACCAAGAGACCTGGAAAGGCGCAGAAATTCAC	647

RESULT 11

US-09-853-625B-1
; Sequence 1, Application US/09853625B
; Patent NO. US20020049304A1
; GENERAL INFORMATION:

APPLICANT: Haellings, Gregg A. and Adams, Mark D.
TITLE OF INVENTION: Human CCN-Like Growth Factor
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARBELL, BYRNE, BAIN, GILFILLIAN,
CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/853,625B
FILING DATE: 14-May-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/053,587
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: MULLINS, J.G.
REGISTRATION NUMBER: 33,073
REFERENCE/DOCKET NUMBER: 325800-442
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 900 BASE PAIRS
TYPE: NUCLEIC ACID
STRANDEDNESS: SINGLE
TOPOLOGY: LINEAR
MOLECULE TYPE: CDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-853-625B-1

Query Match 10.6%; Score 67.2; DB 10; Length 900;
Best Local Similarity 50.9%; Pred. No. 2.9e-09;
Matches 274; Conservative 0; Mismatches 243; Indels 21; Gaps 4;
QY 82 GCCTTCAGAGATGATGACAGAGATGATCCAGAGGCTTGAGAGTACCCCGAGC---CT 138
DB 184 GCTTTAAATGATGACAGAGATGATCTTTATTCATGATGTTAACTCTGTTCCAGCA 243
QY 139 CTTCTGAGAACCAACGACCATGAAACCGGCGAGAGATGAGGACAGACTCC---CAC 195
DB 244 CACCCGAGAGCAACAGCAGTGAATCAAGCCAGAAATGAGGAGGAGCTTTCACTAAC 303
QY 196 CATCCCTATGACGCAAGATGATGTCGAGTACAGCTCCGCGAGCTGCACTACACCCGC 255
DB 304 ACTGAGTGTATGGAACACTCGGTTCAAGTGGTTGCGGAAACTGCTTCACCAAA 363
QY 256 TTCTGACAGAGAGGCGCATGCGGAGCGGCAAGCGGTCACCGAGTTGATGTCGCGG 315
DB 364 TACATCTCTATGAGCCAGTGCACACGATCAGCCCTTGAAAGAGCTGATGTCGCGC 423
QY 316 CAGTGGGCGCCCGCGGCTGTCGCCAACGCCATCGGGCGCT-----GAA 363
DB 424 GAGTGTGTCGCGCCGAGTGTCTCCCTAATGATGAGAGGCTATGAAACAAAGTAC 483
QY 364 TGTGTGGCGCCGAGAGCGGATTTCCGTCATCCCGGATGCGTACCCGCGGAGCGG 423
DB 484 TGGAGCAGAGAGAGCTCCAGAGAGTGGGTGTCTAATACAAACCTGTACCCAGAGA 543
QY 424 GTGAGCTGCTGTGCGCCGCGGAGCGCGCGCTCGCGCAAGGTGCTGTGTGAGCC 483
DB 544 ATCCAGCTGAGTGCAGAAATGAGCAGC---ACAGCAGCTACAAATACAGATGACACT 600
QY 484 TCGTCAAGTGCAGAGCGCTCAGCCGCTTCACAAACAGTGCAGAGCTCAAGAGCTTCGGG 543

DB 601 GCGTCAAGTGCAGAGGATGACACCCGCGAGCAGACAGAGTCCAGTACAACTTTGAGAGC 660
QY 544 CCGAGAGCCGCGCGCGCGAGAGAGGTGCGCAAGCCGCGCGCGCGCGGAGAGCA 601
DB 661 ATGTACCTGCGAAGCAGTCCAGATCAGAGAGCGGAAAGAGCAGCAATCCA 718
RESULT 12
US-09-745-288-92/c
; Sequence 92, Application US/09745288
; Patent No. US20010018058A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, David C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; FILE REFERENCE: 210121.446D1
; CURRENT APPLICATION NUMBER: US/09/745,288
; CURRENT FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 92
; LENGTH: 1692
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-745-288-92

Query Match 10.3%; Score 65.6; DB 10; Length 1692;
Best Local Similarity 50.7%; Pred. No. 8e-09;
Matches 273; Conservative 0; Mismatches 244; Indels 21; Gaps 4;

QY 82 GCCTTCAGAGATGATGACAGAGTATCCAGAGGCTTGAGAGTACCCCGAGC---CT 138
DB 1574 GCTTTAAATGATGACAGAGATGATCTTTATTCATGATGTTAACTCTGTTCCAGCA 1515
QY 139 CTTCTGAGAACCAACGACCATGAAACCGGCGAGAGATGAGGACAGACTCC---CAC 195
DB 1514 CACCCGAGAGCAACAGCAGTGAATCAAGCCAGAAATGAGGAGGAGCTTTCACTAAC 1455
QY 196 CATCCCTATGACGCAAGATGATGTCGAGTACAGCTCCGCGAGCTGCACTACACCCGC 255
DB 1454 ACTGAGTGTATGGAACACTCGGTTCAAGTGGTTGCGGAACTGCTTCACCAAA 1395
QY 256 TTCTGACAGAGAGGCGCATGCGGAGCGGCAAGCGGTCACCGAGTTGATGTCGCGG 315
DB 1394 TACATCTCTATGAGCCAGTGCACACGATCAGCCCTTGAAAGAGCTGATGTCGCGC 1335
QY 316 CAGTGGGCGCCCGCGGCTGTCGCCAACGCCATCGGGCGCT-----GAA 363
DB 1334 GAGTGTGTCGCGCCGAGTGTCTCCCTAATGATGAGAGGCTATGAAACAAAGTAC 1275
QY 364 TGTGTGGCGCCGAGAGCGGATTTCCGTCATCCCGGATGCGTACCCGCGGAGCGG 423
DB 1274 TGGAGCAGAGAGAGCTCCAGAGAGTGGGTGTCTAATACAAACCTGTACCCAGAGA 1215
QY 424 GTGAGCTGCTGTGCGCCGCGGAGCGCGCGCTCGCGCAAGGTGCTGTGTGAGCC 483
DB 1214 ATTCAGCTGATGCAAGATGAGTGCAGC---ACAGCAGCTACAAATACAGTATCTACT 1158
QY 484 TCGTCAAGTGCAGAGCGCTCAGCCGCTTCACAAACAGTGCAGAGCTCAAGAGCTTCGGG 543
DB 1157 GCGTCAAGTGCAGAGAGTGCACCCGCGAGCAGACAGATGCAAGTCAAGTCAAGTGAAGC 1098
QY 544 CCGAGAGCCGCGCGCGCGAGAGAGGTGCAAGCGCGCGCGCGCGCGCGGAGAGCA 601
DB 1097 ATGTACCTGCAAGCAGTCCAGATCAGAGAGCGGAAAGAGCAGCAAAATCCA 1040

RESULT 13
US-09-803-719-2375
; Sequence 2375, Application US/09803719

```

Publication No. US20030044783A1
GENERAL INFORMATION:
APPLICANT: Williams, Lewis T.
APPLICANT: Escobedo, Jaime
APPLICANT: Inniss, Michael A.
APPLICANT: Garcia, Pablo Dominguez
APPLICANT: Sudduth-Klinger, Julie
APPLICANT: Reinhard, Christoph
APPLICANT: Giese, Klaus
APPLICANT: Randazzo, Filippo
APPLICANT: Kennedy, Giulia C.
APPLICANT: Pot, David
APPLICANT: Kassam, Alraf
APPLICANT: Iamson, George
APPLICANT: Drmanac, Radoje
APPLICANT: Crkvenjakov, Radomir
APPLICANT: Dickson, Mark
APPLICANT: Drmanac, Snezana
APPLICANT: Iabat, Ivan
APPLICANT: Leshkowitz, Dena
APPLICANT: Kita, David
APPLICANT: Garcia, Veronica
APPLICANT: Jones, Lee William
APPLICANT: Strache-Crain, Biggit
TITLE OF INVENTION: Human Genes and Gene Products
FILE REFERENCE: 1624, 002
CURRENT APPLICATION NUMBER: US/09/803, 719
CURRENT FILING DATE: 2001-03-09
PRIOR FILING DATE: 2000-03-09
PRIOR FILING DATE: 2000-03-09
NUMBER OF SEQ ID NOS: 2396
SOFTWARE: Fast-Seq for Windows Version 4.0
SEQ ID NO 2375
LENGTH: 373
TYPE: DNA
ORGANISM: Homo sapiens
US-09-803-719-2375

```

Query Match	6.9%	Score	44.2	DB	9	Length	373
Best Local Similarity	53.8%	Pred. No.	0.0045				
Matches	91	Conservative	0	Mismatches	78	Indels	0
						Gaps	0

Qy	396	CATCCCGGATGCGTCAACCGCGGACAGGGGTACAGTGTGAGCCCGGAGGCGGGGCGC	455
Dy	93	CAAGCAGGAGCAGCTACCGCTGACAGGAGCTCAGGGGCTCTGGGGGCGGGGGGGGCGT	152
Dy	456	GCGCTCGCGCMAAGTGTGCTGTGTGGCTCTGTGCAAGTGTGAAGGGCGCTCACCGGCTTCCA	515
Dy	153	GCGCAGGCGCCCGACGGCGCGCTTTGTGCGCCCTTCACAGCTACGCAAGCTACGCTTCCCTTCAA	212
Qy	516	CACCCAGTTCGAGCTCAAGAGACTTCGGGCGGAGACCGCGCGGCGCGAG	564
Dy	213	GAGCAACCCCCGAGCACGGGAGTCTTTGGCTCAGGGATGGGAGCGGGAGG	261

RESULT 14
 US-98-803-719-2377
 Sequence 2377, Application US/09803719
 Publication No. US20030044783A1
 GENERAL INFORMATION:
 APPLICANT: Williams, Lewis T.
 APPLICANT: Escobedo, Jaime
 APPLICANT: Imis, Michael A.
 APPLICANT: Garcia, Pablo Dominguez
 APPLICANT: Sudduth-Ringer, Julie
 APPLICANT: Reinhard, Christoph
 APPLICANT: Giese, Klaus
 APPLICANT: Randazzo, Filippo
 APPLICANT: Kennedy, Giulia C.
 APPLICANT: Pot, David
 APPLICANT: Kaasam, Altaf
 APPLICANT: Lamson, George
 APPLICANT: Drmanac, Radoje

```

1  APPLICANT: Cirkvenjakov, Radomir
2  APPLICANT: Dickson, Mark
3  APPLICANT: Driemac, Snezana
4  APPLICANT: Labat, Ivan
5  APPLICANT: Leschkowitz, Dena
6  APPLICANT: Kita, David
7  APPLICANT: Garcia, Veronica
8  APPLICANT: Jones, Lee William
9  APPLICANT: Stache-Crain, Biggit
10 TITLE OF INVENTION: Human Genes and Gene
11 FILE REFERENCE: 1644.002
12 CURRENT APPLICATION NUMBER: US/09/803,719
13 CURRENT FILING DATE: 2001-03-09
14 PRIOR APPLICATION NUMBER: 60/188,609
15 PRIOR FILING DATE: 2000-03-09
16 NUMBER OF SEQ ID NOS: 2396
17 SOFTWARE: fastSeq for Windows Version 4.0.
18 SEQ ID NO 2377
19 LENGTH: 426
20 TYPE: DNA
21 ORGANISM: Homo sapiens
22 US-09-803-719-2377

```

Query Match	6.9%	Score 44.2;	DB 9;	Length 426;
Best Local Similarity	53.8%;	Pred. No. 0.0045;		
Matches 91; Conservative	0;	Mismatches 78;	Indels 0;	Gaps 0;

QY	396	CAATCCGAGTACCTACCGCGCGAGAGAGGTCAGACTGTGCCCCCGAGAGGCGCGAGCGCC	455
QY	54	CAGACAGACAGACTACCGGCTGCGAGAGACTGACAGGCTCTGGGGGCGCGGCGGAGGAGCT <td>113</td>	113
Db	456	GGCGTCGCGAGAGTGCCTCTTGCTGAGCGCTGATGACAGTCGACAGCGGCTCACCGGCTTCCA	515
QY	114	GGCAGAGCCCCGAGCGCGCTTTGAGCCCCCTCACAGACTGAGACTACGCTTCCCTTCTTCA	173
Db			

QY	516	CAACCGAGTCGAGCTCAAGGACTTCGGGCCCGAGACCGCGGGCCGAC	564
Db	174	GAGCAACCCCGACGACCGGGAGTCTTGGCCAGCGGATGGGGAGCGGGGAG	222

```

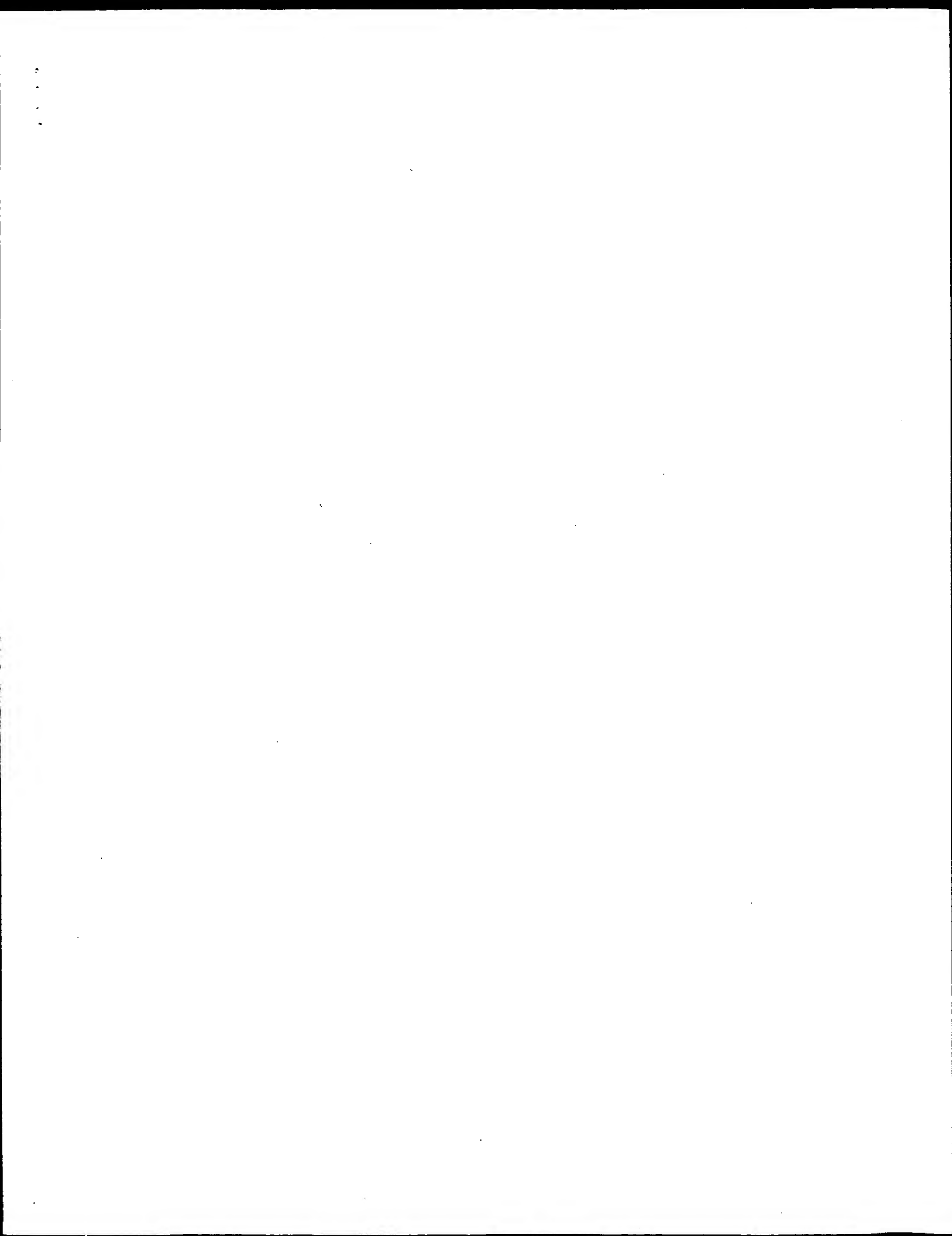
RESULT 15
US-09-822-830A-91
? Sequence 91, Application US/09822830A
? Patent No. US20020142952A1
? GENERAL INFORMATION:
? APPLICANT: Genetics Institute, Inc.
? APPLICANT: Mong, Gordon G.
? APPLICANT: Clark, Hilary
? APPLICANT: Fecthel, Kim
? APPLICANT: Agostino, Michael J.
? APPLICANT: Howes, Steven H.
? APPLICANT: Resnick, Richard J.
? APPLICANT: Gulakota, Kamalakhar
? APPLICANT: Graham, James R.
? TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
? FILE REFERENCE: GIN 6402
? CURRENT APPLICATION NUMBER: US/09/822,830A
? CURRENT FILING DATE: 2001-03-29
? PRIOR APPLICATION NUMBER: 60/125,604
? PRIOR FILING DATE: 2000-04-06
? NUMBER OF SEQ ID NOS: 631
? SOFTWARE: Patentin Ver. 2.0
? SEQ ID NO 91
? LENGTH: 1951
? TYPE: DNA
? ORGANISM: Homo sapiens
US-09-822-830A-91

```

Query	396	CATCCCGGATGCGCTACCCGCGCAGCGGAGTGCCTGTCGCCCGGGGGGCGCGGCC	455
Query Match	6.9%	Score 44.2; DB 10; Length 1951;	
Best Local Similarity	53.8%;	Pred. No. 0.0045;	
Matches	91; Conservative	0; Mismatches 78; Indels	0; Gaps

Db 918 CAAACAGAGAGAGCTACGCTGCAAGGACTGCAAGGCTCTGGGGCCCGGGCGGGGGCT 977
 Qy 456 GCGCTCGGCAAGGTGGCTGTGAGCTGTGCAAGTGCAGAGGCTCAACCGCTTCA 515
 Db 978 GCGCAGGCCCCCAAGCCCTTTGTGCCCCCTCACGACTGCAAGCTACGCTTCCCTTCAA 1037
 Qy 516 CAAACAGTGGAGCTCAAGACTTGGGGCCGAGACCGCGCGCGCAG 564
 Db 1038 GAGCAACCCCAAGCAGGAGTCTTGGCCAGGATGGGAGCGGGAG 1086

Search completed: March 29, 2003, 01:33:24
 Job time : 79.1376 secs



PF 29-MAY-2001; 2001WO-US17478.
 XX
 PR 01-JUN-2000; 2000US-208550P.
 PR 04-AUG-2000; 2000US-223542P.
 XX
 PA (AMGE-) AMGEN INC.
 PI
 PI Paszy CJ, Gao Y;
 XX
 DR WPI: 2002-114325/15.
 DR P-PSDB; ABB07208, ABB07210.
 XX
 PT New human and mouse cysteine-knot polypeptide designated as Cloaked-2,
 PT for treating or preventing kidney, heart (e.g. myocardial infarction)
 PT or liver (e.g. hepatitis) diseases -
 XX
 PS Claim 1; Fig 2; 170pp; English.
 XX
 CC The invention relates to polypeptides comprising a cysteine knot motif
 CC and designated as Cloaked-2, derived from human and mouse. The cloaked-2
 CC polypeptides can be expressed by standard recombinant methodology. The
 CC cloaked-2 polynucleotides are useful in gene therapy and antisense
 CC therapy. The cloaked-2 polypeptides and polynucleotides are useful for
 CC treating, preventing, ameliorating or detecting diseases and disorders of
 CC the kidney (e.g. anemia, hypertension or low blood pressure), heart (e.g.
 CC cardiac hypertrophy, congestive heart failure, myocardial infarction,
 CC arrhythmias, atherosclerosis, hypertension or low blood pressure),
 CC skeletal muscle (e.g. muscular dystrophy or cachexia), placenta (e.g.
 CC congenital abnormalities or miscarriage), liver (e.g. hepatitis or
 CC cirrhosis), pancreas (e.g. diabetes or pancreatitis), thyroid (e.g.
 CC Grave's disease or myxedema) or adrenal cortex (e.g. Cushing's disease
 CC or Addison's disease), homeostasis or metabolic diseases (e.g. obesity,
 CC cancer or myopathies), infections, or autoimmune diseases. Selective
 CC binding agents may be used to modulate the biological activities of
 CC cloaked-2 polypeptides or to detect Cloaked-2 polypeptide levels in a
 CC sample. Transgenic non-human animals are useful for drug candidate
 CC screening. The present sequence represents a cDNA encoding the mouse
 CC cloaked-2 polypeptide.
 CC
 XX
 SQ Sequence 636 BP; 114 A; 224 C; 207 G; 91 T; 0 other;
 Query Match 100.0%; Score 636; DB 24; Length 636;
 Best Local Similarity 100.0%; Pred. No. 8.1e-124;
 Matches 636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGAGCCCTCTAGAGCCCGGCTGATGCTGCTAGTGGACGGCTTCTGTGCT 60
 DB 1 ATGAGCCCTCTAGAGCCCGGCTGATGCTGCTAGTGGACGGCTTCTGTGCT 60
 QY 61 GTGAGGGCCAGGGGTGGCAAGCTTCTGAGATGATCCAGAGTCTATCCAGGGCTT 120
 DB 61 GTGAGGGCCAGGGGTGGCAAGCTTCTGAGATGATCCAGAGTCTATCCAGGGCTT 120
 QY 121 GGAAGTACCCCGAGCCCTCTCTGAGAACCAACCAACATGAACCGGGCGAGATGGA 180
 DB 121 GGAAGTACCCCGAGCCCTCTCTGAGAACCAACCAACATGAACCGGGCGAGATGGA 180
 QY 181 GGCAGACCTCCCAACCTCTATGAGGCCCAAGATGTCCTCGATCAGTCCGGGAG 240
 DB 181 GGCAGACCTCCCAACCTCTATGAGGCCCAAGATGTCCTCGATCAGTCCGGGAG 240
 QY 241 CTGACATACACCCGCTTCTGAGACAGAGCCCATGCGCAGCGCAAGCGGTCACCGAG 300
 DB 241 CTGACATACACCCGCTTCTGAGACAGAGCCCATGCGCAGCGCAAGCGGTCACCGAG 300
 QY 301 TTGCTGTCTCCGCGCAGTGGCGCGCGCGCTGCTGCTCCCAAGCCATCGGGCGGTG 360
 DB 301 TTGCTGTCTCCGCGCAGTGGCGCGCGCGCTGCTGCTCCCAAGCCATCGGGCGGTG 360
 QY 361 AAGGTGAGCGCCCGAAGCGAGCCGATTTCCGCTGACATCCCGATGCTACCGGCGCAG 420
 DB 361 AAGGTGAGCGCCCGAAGCGAGCCGATTTCCGCTGACATCCCGATGCTACCGGCGCAG 420

QY 421 CGAGTGCAGCTCTGTGCCCGGGGGCGGCGCGCGCTGCGCCAAAGTGCCTGTGAGT 480
 DB 421 CGAGTGCAGCTCTGTGCCCGGGGGCGGCGCGCGCTGCGCCAAAGTGCCTGTGAGT 480
 QY 481 GCCTGTGTAATGCAACCGCTCTACCGGCTTCCACAAACGATGGAGCTCAAGACTTC 540
 DB 481 GCCTGTGTAATGCAACCGCTCTACCGGCTTCCACAAACGATGGAGCTCAAGACTTC 540
 QY 541 GGGCCGAGAGCGCGCGCGCGCAGAAAGGATGCGAAGCGCGCGCGCGCGCGGAGCC 600
 DB 541 GGGCCGAGAGCGCGCGCGCGCAGAAAGGATGCGAAGCGCGCGCGCGCGGAGCC 600
 QY 601 AAAGCCAAACGAGCGGAGCTGAGAAACGCTACTAG 636
 DB 601 AAAGCCAAACGAGCGGAGCTGAGAAACGCTACTAG 636
 RESULT 2
 AAA29058
 ID AAA29058 standard; cDNA; 638 BP.
 XX
 AC AAA29058;
 XX
 DT 12-SEP-2000 (first entry)
 XX
 DE Murine TGF-beta binding protein (BBER) cDNA.
 XX
 KW osteoparh; transforming growth factor-beta; TGF-beta; binding protein;
 KW BBER; gene therapy; antisense therapy; fracture; bone mineralization; ss.
 XX
 OS Mus musculus.
 XX
 FH Key location/Qualifiers
 FT CDS 1..636
 FT /tag= a
 FT /product= TGF-beta_binding_protein
 XX
 PN WO200032773-A1.
 XX
 PD 08-JUN-2000.
 XX
 PF 24-NOV-1999; 99WO-US27990.
 XX
 PR 27-NOV-1998; 98US-0110283.
 XX
 PA (DARW-) DARWIN DISCOVERY LTD.
 XX
 PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;
 PI Van Ness J, Winkler Dg;
 XX
 DR WPI: 2000-412321/35.
 PT P-PSDB; AAY96432.
 PT
 PT Nucleic acids (I) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures
 XX
 PS Claim 1; Page 123; 162pp; English.
 XX
 CC This cDNA encodes a murine transforming growth factor-beta (TGF-beta)
 CC binding protein designated mBER. The cDNA and protein may be used for
 CC prevention, treatment and diagnosis of diseases associated with
 CC inappropriate BBER expression. For example, they may be used to treat
 CC disorders associated with decreased TGF-beta BP expression. The cDNA or
 CC vectors may be administered to treat diseases by rectifying mutations or
 CC deletions in a patient's genome that affect the activity of BBER by
 CC expressing inactive proteins or to supplement the patient's own production
 CC of BBER polypeptides. The nucleic acids may be used for recombinant
 CC production of BBER, gene therapy, antisense therapy, as probes for
 CC diagnostic assays and for functional studies. BBER may be used to raise
 CC antibodies and for identification of BBER modulators. BBER antagonists
 CC may be used to increase bone mineral content for the treatment of
 CC disorders such as osteopenia, osteoporosis, fractures and other disorders

CC associated with low mineral content.

XX Sequence 638 BP; 114 A; 224 C; 209 G; 91 T; 0 other;

Query Match 99.7%; Score 634.4; DB 21; Length 638;
Best Local Similarity 99.8%; Pred. No. 1.8e-123;
Matches 635; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ATGACCCCTCCTAGCCCTGCTCATCTGCTCTTGTGACGCTGCTTCTGCT 60
   |||
Db 1 ATGACCCCTCCTAGCCCTGCTCATCTGCTCTTGTGACGCTGCTTCTGCT 60
QY 61 GTGAGGCGCGGAGGCTTCAAGATGATCCACAGAGTCAATCCAGGCTT 120
   |||
Db 61 GTGAGGCGCGGAGGCTTCAAGATGATCCACAGAGTCAATCCAGGCTT 120
QY 121 GGAAGTATCCCGACGCTCTCTTGAGAACACACAGACCTGAGGAAATGGA 180
   |||
Db 121 GGAAGTATCCCGACGCTCTCTTGAGAACACACAGACCTGAGGAAATGGA 180
QY 181 GGCAGACTCCCGACCATCCCTATGACGCGCAAGATGTGTCAGTACAGCTGCGCGAG 240
   |||
Db 181 GGCAGACTCCCGACCATCCCTATGACGCGCAAGATGTGTCAGTACAGCTGCGCGAG 240
QY 241 CTGCACTACACCCGCTTCTGACAGACGCGCCATGCGCGACCGCAAGCCGCTCACCGAG 300
   |||
Db 241 CTGCACTACACCCGCTTCTGACAGACGCGCCATGCGCGACCGCAAGCCGCTCACCGAG 300
QY 301 TTGGTGTCTCCGCGGAGGCGCGCGCGCGCTGCTGCGCAAGCGCTGCGCGCGCTG 360
   |||
Db 301 TTGGTGTCTCCGCGGAGGCGCGCGCGCGCTGCTGCGCAAGCGCTGCGCGCGCTG 360
QY 361 AAGTGTGCGCGCGCGAGACCGGATTTCCGCTGATCCCGGATGCTAACCGCGCGAG 420
   |||
Db 361 AAGTGTGCGCGCGAGACCGGATTTCCGCTGATCCCGGATGCTAACCGCGCGAG 420
QY 421 CGGGTCAAGTGTGTGCTCCCGGGGGCGCGCGCGCTGCGCAAGTGTGCTTGTG 480
   |||
Db 421 CGGGTCAAGTGTGTGCTCCCGGGGGCGCGCGCGCTGCGCAAGTGTGCTTGTG 480
QY 481 GCTCTGTCAGAGTGAAGGCGCTCACCGGCTTCAACACAGCGGAGCTCAAGGACTTC 540
   |||
Db 481 GCTCTGTCAGAGTGAAGGCGCTCACCGGCTTCAACACAGCGGAGCTCAAGGACTTC 540
QY 541 GGGCGGAGACCGCGCGCGCGCGAGAGGTGCGAAGCGCGCGCGCGCGCGGAGGC 600
   |||
Db 541 GGGCGGAGACCGCGCGCGCGCGAGAGGTGCGAAGCGCGCGCGCGCGGAGGC 600
QY 601 AAGCCACACGCGGAGGCTGAGAAAGCTTACTAG 636
   |||
Db 601 AAGCCACACGCGGAGGCTGAGAAAGCTTACTAG 636

```

RESULT 3

AA29059
ID AAA29059 standard; cDNA; 674 BP.

```

XX AAA29059;
XX
XX 12-SEP-2000 (first entry)
XX
XX Rat TGF-beta binding protein (BEER) cDNA.
XX
XX osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
XX BEER; gene therapy; antisense therapy; fracture; bone mineralization; ss.
XX
XX Rattus norvegicus.
XX
XX
XX Key Location/Qualifiers
XX CDS 33..674
XX FT /tag= a
XX FT /product= TGF-beta_binding_protein
XX

```

```

PN WO200032773-A1.
XX
XX 08-JUN-2000.
XX
XX 24-NOV-1999; 99WO-US27990.
XX
XX 27-NOV-1998; 98US-0110283.
XX
XX (DARWIN) DARWIN DISCOVERY LTD.
XX
XX Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW,
XX Van Ness J, Winkler DG;
XX
XX WPI; 2000-412321/35.
XX
XX P-P9DB; AAY96433.
XX
XX
XX Nucleic acids (1) encoding a transforming growth factor beta binding
XX protein, useful for identifying agents for treating osteopenia,
XX osteoporosis and fractures
XX
XX
XX Claim 1; Page 125; 162pp; English.
XX
XX
XX This cDNA encodes a rat transforming growth factor-beta (TGF-beta)
XX binding protein designated rBEER. The cDNA and protein may be used for
XX prevention, treatment and diagnosis of diseases associated with
XX inappropriate BEER expression. For example, they may be used to treat
XX disorders associated with decreased TGF-beta BP expression. The cDNA or
XX vectors may be administered to treat diseases by rectifying mutations or
XX deletions in a patient's genome that affect the activity of BEER by
XX expressing inactive proteins or to supplement the patient's own production
XX of BEER polypeptides. The nucleic acids may be used for recombinant
XX production of BEER, gene therapy, antisense therapy, as probes for
XX diagnostic assays and for functional studies. BEER may be used to raise
XX antibodies and for identification of BEER modulators. BEER antagonists
XX may be used to increase bone mineral content for the treatment of
XX disorders such as osteopenia, osteoporosis, fractures and other disorders
XX associated with low mineral content.
XX
XX
XX Sequence 674 BP; 128 A; 240 C; 208 G; 98 T; 0 other;
XX

```

Query Match 87.7%; Score 557.6; DB 21; Length 674;
Best Local Similarity 93.0%; Pred. No. 1.8e-107;
Matches 597; Conservative 0; Mismatches 39; Indels 6; Gaps 1;

```

QY 1 ATGACCCCTCCTAGCCCTGCTCATCTGCTCTTGTGACGCTGCTTCTGCT 60
   |||
Db 33 ATGACCTCTCACTAGCCCTTGTGCTGCTGCTGCTTGTATGACGCTTGTGCT 92
QY 61 GTGAGGCGCGGAGGCTTCAAGATGATCCACAGAGTCAATCCAGGCTT 120
   |||
Db 93 GTGAGGCGCGGAGGCTTCAAGATGATCCACAGAGTCAATCCAGGCTT 152
QY 121 GGAAGTATCCCGACGCTCTT-----CTGAGAACACACGACCATGAAACCGGCGAG 174
   |||
Db 153 AGAGAGTATCCAGAGCTCTTCAAGAACTAGAGAACACGACCATGAAACCGGCGAG 212
QY 175 AATGAGGCGAGACCTCCCGACCATCTCTATGACGCGCAAGATGTGTCGAGTACAGCTGC 234
   |||
Db 213 AATGAGGCGAGACCTCCCGACCATCTCTATGACACACAAAGACGTGTCCAGTACAGCTGC 272
QY 235 CGCGAGCTGCACTACACCGGCTTCTGACAGAGGCGCGCATGCGCGCAAGCGCAACCGGCTC 294
   |||
Db 273 CGCGAGCTGCACTACACCGGCTTCTGACAGAGGCGCGCATGCGCGCAAGCGCAACCGGCTC 332
QY 295 ACCGAGTGTGTGCTCCGCGAGTGGCGCGCGCGCGCGCTGCGCGCAAGCGCATGCGG 354
   |||
Db 333 ACCGAGTGTGTGCTCCGCGAGTGGCGCGCGCGCGCGCTGCGCGCAAGCGCATGCGG 392
QY 355 CGCGTGAAGTGTGGCGCGCGAGACGAGATTTCCGCTGATCCCGATCGCTACCGC 414
   |||
Db 393 CGCGTGAAGTGTGGCGCGCGAGACGAGCTTCCGCTGATCCCGATCGCTACCGC 452
QY 415 GCGCAGCGGGTGAAGTGTGCTGCTCCCGGGGGCGCGCGCGCTGCGCGCAAGTGTGCT 474

```


XX MO200055193-A2.
 XX 21-SEP-2000.
 XX
 XX 02-MAR-2000; 2000WO-US05537.
 XX
 XX 12-MAR-1999; 99US-0124118.
 XX (REG-) REGENERON PHARM INC.
 XX
 XX Economides AN;
 XX
 XX WPI; 2000-638179/61.
 XX P-PSDB; AAB6106.
 XX
 XX Novel isolated, human DNA/Cerberus related protein 6 which include
 XX natural homologue, and polypeptides comprising DCR6 domain and nucleic
 XX acids encoding the proteins which are useful as probes and primers -
 XX
 XX Claim 2; Fig 3; 40pp; English.
 XX
 XX The present sequence comprises exons 1 and 4 of the human
 XX DAN/Cerberus-related protein 6 (hDCR6) coding sequence. It was isolated
 XX from a human kidney cDNA library containing exons 1 and 4 of the
 XX sequence. hDCR6 is closely related to the DAN and DCR5 proteins, both of
 XX which act as antagonists of morphogenic proteins such as BMP. It is
 XX possible that the hDCR6 gene and protein can be used as immunogens,
 XX modulators of cell function, growth and differentiation, to reduce
 XX undesirable bone formation, to identify DCR6 binding agents, in
 XX diagnosis, and in gene therapy.
 XX
 XX Sequence 642 BP; 115 A; 230 C; 211 G; 86 T; 0 other;
 XX
 XX Query Match 78.9%; Score 501.6; DB 21; Length 642;
 XX Best Local Similarity 87.5%; Pred. No. 8; 7e-96;
 XX Matches 562; Conservative 0; Mismatches 74; Indels 6; Gaps 1;
 XX
 XX 1 ATGCAGCCCTCACTAGCCCGCTGCTCATCTGCTACTTGTGACCGCTGCTGCTGCT 60
 XX 1 ATGCAGCTCCCACTGGCCCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
 XX
 XX 61 GTGGAGGGCCAGGGGGTGGCAGGCTTCAGGAATGATGCCACAGAGGTCAATCCAGGGCTT 120
 XX 61 GTGGAGGGCCAGGGGGTGGCAGGCTTCAGGAATGATGCCACAGAGGTCAATCCAGGGCTT 120
 XX
 XX 121 GGAGAGTACCCCGAGGCTCTCTCC-----TGAGACACACAGCCATGAAACCGGGGGAG 174
 XX 121 GGAGAGTACCCCGAGGCTCTCTCC-----TGAGACACACAGCCATGAAACCGGGGGAG 174
 XX
 XX 175 AATGAGGCGAGACCTCCCAACATCCCTATGACGCCAAGATGTCTGAGTACAGCTGC 234
 XX 181 AATGAGGCGAGGCTCCCAACATCCCTATGAGACGCCAAGATGTCTGAGTACAGCTGC 240
 XX
 XX 235 CGGAGCTGCACTACACCCGCTTCTGACAGACGCGCCATGCCGACGCGCAAGCGGCTC 294
 XX 241 CGGAGCTGCACTACACCCGCTTCTGACAGACGCGCCATGCCGACGCGCGCAAGCGGCTC 300
 XX
 XX 295 ACCGAGTGGTGTGCTCCGGCCAGTGGGGCCGCGCGCGCTGTGCGCCAAACGCAATCGG 354
 XX 301 ACCGAGTGGTGTGCTCCGGCCAGTGGGGCCGCGCGCGCTGTGCGCCAAACGCAATCGG 360
 XX
 XX 355 CGGTAAGTGTGTGCGCCCGAAGCGAGATTTCCGCTGATCCGCGATCGCTACCGC 414
 XX 361 CGGTAAGTGTGTGCGCCCGAAGCGAGATTTCCGCTGATCCGCGATCGCTACCGC 420
 XX
 XX 415 GCGCAGCGGTGCACTGCTGTGCTCCCGGGGGCGCGCGCGCTGCGCAAGTGCGCT 474
 XX 421 GCGCAGCGGTGCACTGCTGTGCTCCCGGGGGCGCGCGCGCTGCGCAAGTGCGCGC 480
 XX
 XX 475 CTGGGAGCGCTCGCAAGTGCAGAGGCGCTCAACCGCTTCACAAACAGTCGAGCTCAAG 534
 XX 481 CTGGGAGCGCTCGCAAGTGCAGAGGCGCTCAACCGCTTCACAAACAGTCGAGCTCAAG 540

QY 535 GACTTCGGGCGGAGACCGCGCGCGAGAAAGGTGCGCAAGCCGCGCGCGCGCGCG 594
 DB 541 GACTTCGGGCGGAGACCGCGCGCGAGAAAGGTGCGCAAGCCGCGCGCGCGCGCG 600
 QY 595 GGAGCCAAAGCCACACGAGCGGAGCTGAGAAACGCTACTAG 636
 DB 601 AGCGCCAAAGCCACACGAGCGGAGCTGAGAAACGCTACTAG 642
 RESULT 6
 ABA94293
 ID ABA94293 standard; cDNA; 759 BP.
 AC ABA94293;
 XX
 XX 26-MAR-2002 (first entry)
 XX
 XX Human cloaked-2 polypeptide encoding cDNA.
 DE
 XX Cloaked-2; cysteine knot motif; nephroretropic; cardiac; immunomodulator;
 KW hepatotropic; antiinflammatory; antithyroid; cytostatic; neuroprotective;
 KW antianemic; hypotensive; antiarrhythmic; antihypertensive; muscular;
 KW antidiabetic; anorectic; gene therapy; cell therapy; antisense therapy;
 KW human; ss.
 XX
 XX Homo sapiens.
 PH Key Location/Qualifiers
 FT CDS 37..678
 FT /*tag= a
 FT /product= "cloaked-2 polypeptide"
 FT sig_peptide 37..105
 FT /*tag= b
 FT mat_peptide 106..759
 FT /*tag= c
 FT /note= "see ABB07207"
 XX
 XX WO200192308-A2.
 XX
 XX 06-DEC-2001.
 XX
 XX 29-MAY-2001; 2001WO-US17478.
 XX
 XX 01-JUN-2000; 2000US-208550P.
 XX
 XX 04-AUG-2000; 2000US-223542P.
 XX
 XX (AMGE-) AMGEN INC.
 XX
 XX Paeszy CJ, Gao Y;
 XX
 XX WPI; 2002-114325/15.
 XX
 XX P-PSDB; ABB07207, ABB07209.
 XX
 XX
 XX Claim 1; Fig 1; 170pp; English.
 XX
 XX The invention relates to polypeptides comprising a cysteine knot motif
 XX and designated as Cloaked-2, derived from human and mouse. The cloaked-2
 XX polypeptides can be expressed by standard recombinant methodology. The
 XX cloaked-2 polynucleotides are useful in gene therapy and antisense
 XX therapy. The cloaked-2 polypeptides and polynucleotides are useful for
 XX treating, preventing, ameliorating or detecting diseases and disorders of
 XX the kidney (e.g. anemia, hypertension or low blood pressure), heart (e.g.
 XX cardiac hypertrophy, congestive heart failure, myocardial infarction,
 XX arrhythmias, atherosclerosis, hypertension or low blood pressure),
 XX skeletal muscle (e.g. muscular dystrophy or cachexia), placenta (e.g.
 XX congenital abnormalities or miscarriage), liver (e.g. hepatitis or
 XX cirrhosis), pancreas (e.g. diabetes or pancreatitis), thyroid (e.g.
 XX Grave's disease or myxedema) or adrenal cortex (e.g. Cushing's disease

or Addison's disease), homeostasis or metabolic diseases (e.g. obesity, cancer or myopathies), infections, or autoimmune diseases. Selective binding agents may be used to modulate the biological activities of Cloaked-2 polypeptides or to detect Cloaked-2 polypeptide levels in a sample. Transgenic non-human animals are useful for drug candidate screening. The present sequence represents a cDNA encoding the human cloaked-2 polypeptide.

Sequence 759 BP; 125 A; 282 C; 244 G; 108 T; 0 other;

Query Match 78.9%; Score 501.6; DB 24; Length 759;
Best Local Similarity 87.5%; Pred. No. 8.7e-96;
Matches 562; Conservative 0; Mismatches 74; Indels 6; Gaps 1;

```

QY 1 ATGAGCCCTCCTAGACCCGCTGCTCATCTGCTACTTGTGACCGCTGCTTGTGCT 60
DB 37 ATGAGCTCCACATGCGCTGCTGTCTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 96
QY 61 GTGAGAGGCGCAGGGGTGGCAAGCCTTCAGAGATGATGCCACAGAGGTCATCCGAGGCTT 120
DB 97 GTGAGAGGCGCAGGGGTGGCAAGCCTTCAGAGATGATGCCACAGAGGTCATCCGAGGCTT 156
QY 121 GGAGAGTACCCCGAGCCTCTCC-----TGAGAACACCAAGACCATGAAACCGGCGGAG 174
DB 157 GGAGAGTACCCCGAGCCTCTCCACCGAGAGCTGAGAACCAAGACCATGAAACCGGCGGAG 216
QY 175 AATGAGAGGAGACCTCCCAACATCCCTATGAGCGCCAAAGATGTCCTGAGTACAGCTGC 234
DB 217 AATGAGAGGAGACCTCCCAACATCCCTATGAGCGCCAAAGATGTCCTGAGTACAGCTGC 276
QY 235 CGGAGCTGACATACACCCGCTTCTGACAGAGGCGCCATGCGCAGCGCAAGCGGCTC 294
DB 277 CGGAGCTGACATACACCCGCTTCTGACAGAGGCGCCATGCGCAGCGCAAGCGGCTC 336
QY 295 ACCGAGTGTGTGCTCTCCGCGAGTGGCGCCCGCGCGCTGCTGCCAACGCGCATCGG 354
DB 337 ACCGAGTGTGTGCTCTCCGCGAGTGGCGCCCGCGCGCTGCTGCCAACGCGCATCGG 396
QY 355 CGGCTGAAGTGTGGCGCCCGCAACGAGATTTCCGCTGCAATCCCGGATCGCTACCG 414
DB 397 CGGCTGAAGTGTGGCGCCCGCAACGAGATTTCCGCTGCAATCCCGGATCGCTACCG 456
QY 415 GCGCAGCGGCTGACGCTGTGCTGCGCGGCGCGCGCGCTGCGCAGAGTGCCT 474
DB 457 GCGCAGCGGCTGACGCTGTGCTGCGCGGCGCGCGCGCTGCGCAGAGTGCCT 516
QY 475 CTGCTGCTCTGCTGCAAGTGAAGCGCTTCAACCGCTTCCACCAACAGTGCAGTCAAG 534
DB 517 CTGCTGCTCTGCTGCAAGTGAAGCGCTTCAACCGCTTCCACCAACAGTGCAGTCAAG 576
QY 535 GACTTCGGGCGGAGACCGCGCGCGCGCGCGAGAGTGTGCAAGCGCGCGCGCGCG 594
DB 577 GACTTCGGGCGGAGACCGCGCGCGCGCGCGAGAGTGTGCAAGCGCGCGCGCGCG 636
QY 595 GAGGCCCAAGCCAGCAGCGCGAGTGTGAGAACCGCTTACTAG 636
DB 637 AGCGCCAAAGCCAGCAGCGCGAGTGTGAGAACCGCTTACTAG 678

```

RESULT 7
AAD27576
ID AAD27576 standard; cDNA; 2271 BP.

AC AAD27576;

DT 18-APR-2002 (first entry)

DE Human osteolevin cDNA.

XX Human, osteolevin; osteopathic; cytostatic; bone formation; osteoporosis;
KW Van Buchem-sclerosteosis disease; sclerosteosis; transgenic animal;
KM Paget's disease; chromosome 17; ss.
XX

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 13..654

FT sig_peptide /tag= a

FT mat_peptide /product= "Human osteolevin protein"

FT /tag= b

FT /tag= c

FT /product= "Mature human osteolevin protein"

PD 27-DEC-2001.

PF 15-JUN-2001; 2001WO-EP06795.

PR 19-JUN-2000; 2000EP-0112867.

PA (HOPE) HOFMANN LA ROCHE & CO AG F.

PI (UNIV-) UNIV INSTELLING ANTWERPEN UTA.

PI Vickerly BH;

PI Balemans W, Ebelling M, Foerzler D, Patel N, Van Hul W;

DR MPI; 2002-139789/18.

DR P-PSDB; AAE17089.

PT Novel genetic polymorphisms in the Van Buchem-sclerosteosis disease

PT region that are associated with abnormal bone formation useful for

PT diagnosis and assessment of osteoporosis or sclerosteosis in humans

PS Claim 10; Fig 1B; 70pp; English.

XX The invention relates to a nucleic acid encoding osteolevin region

CC polymorphisms. The invention also relates to genetic polymorphisms in

CC the Van Buchem-sclerosteosis disease region that are associated with

CC disorders resulting in either net excess bone formation or insufficient

CC bone formation in humans. Osteolevin DNA is useful for screening for

CC osteolevin polymorphisms associated with abnormal bone formation in a

CC subject and for the presence of a heritably linked form of abnormal bone

CC formation in a subject, by determining the presence of a polymorphism in

CC the osteolevin nucleic acid sequence obtained from the subject.

CC Osteolevin protein is useful for treating diseases associated with

CC abnormal bone formation, such as sclerosteosis, Van Buchem's disease and

CC Paget's disease. Nucleic acids which encode genes in the osteolevin

CC region or their modified forms can also be used to generate either

CC transgenic animals or knockout animals which are useful in the screening

CC and development of therapeutically useful reagents. Osteolevin proteins

CC are useful in pharmacological characterisation of novel modulators of the

CC activity of protein and protein complexes. Human osteolevin gene is

CC located on chromosome 17. The present sequence is human osteolevin cDNA.

XX Sequence 2271 BP; 573 A; 623 C; 599 G; 476 T; 0 other;

Query Match 78.9%; Score 501.6; DB 24; Length 2271;
Best Local Similarity 87.5%; Pred. No. 9.2e-96;
Matches 562; Conservative 0; Mismatches 74; Indels 6; Gaps 1;

```

QY 1 ATGAGCCCTCCTAGACCCGCTGCTCATCTGCTACTTGTGACGCTGCTTGTGCT 60
DB 13 ATGAGCTCCACATGCGCTGCTGTCTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 72
QY 61 GTGAGAGGCGCAGGGGTGGCAAGCCTTCAGAGATGATGCCACAGAGGTCATCCGAGGCTT 120
DB 73 GTGAGAGGCGCAGGGGTGGCAAGCCTTCAGAGATGATGCCACAGAGGTCATCCGAGGCTT 132
QY 121 GGAGAGTACCCCGAGCCTCTCC-----TGAGAACACCAAGACCATGAAACCGGCGGAG 174
DB 133 GGAGAGTACCCCGAGCCTCTCCACCGAGCTGAGAACCAAGACCATGAAACCGGCGGAG 192
QY 175 AATGAGAGGAGACCTCCCAACATCCCTATGAGCGCCAAAGATGTCTCGAGTACAGCTGC 234

```

```

Db 193 AACGAGAGGCGGCGCTCCACACCCCTTTGAGACCAAGACGTGTCCGAGTACAGCTGC 252
Qy 235 CGCGAGCTGCACTACACCGCGCTTCTGACAGAGCCCATATGCGGACGCGCAAGCGGTC 294
Db 253 CGCGAGCTGCACTTCAACCGCGCTTCTGACAGAGCCCATATGCGGACGCGCAAGCGGTC 312
Qy 295 ACCGAGTTGTGTCTCCGCGCAAGTGGCGGCGCGCGGCTGCTGCGCAAGCGGTCGCG 354
Db 313 ACCGAGTTGTGTCTCCGCGCAAGTGGCGGCGCGCGGCTGCTGCGCAAGCGGTCGCG 372
Qy 355 CGCGTGAAGTGTGTGCGCGCGCAAGGACCGGATTTCCGCTGACATCCCGATCCGCAAC 414
Db 373 CGCGCAAGTGTGTGCGCGCAAGTGGCGGCGCGGCTGCTGCGCAAGCGGTCGCG 432
Qy 415 GCGCAGCGGCTGCACTGCTGTGCGCGCGGCGCGCGGCGCGGCTGCTGCGCAAGTGTG 474
Db 433 GCGCAGCGGCTGCACTGCTGTGCGCGCGGCGCGGCGCGGCGCGGCTGCTGCGCAAGTGTG 492
Qy 475 CTGCTGCGCTTCTGTCAGAGTGCAGAGCGGCTTCAACCGCTTCAACAGTGCAGTCAAG 534
Db 493 CTGCTGCGCTTCTGTCAGAGTGCAGAGCGGCTTCAACCGCTTCAACAGTGCAGTCAAG 552
Qy 535 GACTTTCGCGCGCGGAGACCGCGCGCGCGCGCGCAAGGCTGCGCAAGCGGCGCGCGCG 594
Db 553 GACTTTCGCGCGCGGAGACCGCGCGCGCGCGCGCAAGGCTGCGCAAGCGGCGCGCGCG 612
Qy 595 GGAGCCAAAGCCAAACGAGCGGAGCTGAGAGAGCGCTTCTTAC 636
Db 613 AGCGCCAAAGCCAAACGAGCGGAGCTGAGAGAGCGCTTCTTAC 654

RESULT 8
AAA29055
ID AAA29055 standard; cDNA; 2301 BP.
XX
AC AAA29055;
XX
DT 12-SBP-2000 (first entry)
XX
DE Human TGF-beta binding protein (BEBR) cDNA.
XX
KM osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
KM BEER; chromosome 17q12-21; gene therapy; antisense therapy; fracture;
XX bone mineralization; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 48..689
FT /tag=a
FT /product=TGF-beta_binding_protein
XX
PN MO200032773-A1.
XX
PD 08-JUN-2000.
XX
PF 24-NOV-1999; 99WO-US27990.
XX
PR 27-NOV-1998; 98US-0110283.
XX
PA (DARW-) DARWIN DISCOVERY LTD.
XX
PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepker BW;
XX Van Ness J, Winkler DG;
XX
DR MPI: 2000-412321/35.
XX P-PSDB; AA196429.
XX
PT Nucleic acids (1) encoding a transforming growth factor beta binding
PT protein, useful for identifying agents for treating osteopenia,
PT osteoporosis and fractures
XX

```

```

PS Claim 1; Page 114-115; 162pp; English.
XX
CC This cDNA encodes a human transforming growth factor-beta (TGF-beta)
CC binding protein designated BEBR. The BEBR gene has been localized
CC to the chromosome 17q12-21. The cDNA and protein may be used for
CC prevention, treatment and diagnosis of diseases associated with
CC inappropriate BEBR expression. For example, they may be used to treat
CC disorders associated with decreased TGF-beta BP expression. The cDNA or
CC vectors may be administered to treat diseases by rectifying mutations or
CC deletions in a patient's genome that affect the activity of BEBR by
CC expressing inactive proteins or to supplement the patients own production
CC of BEBR polypeptides. The nucleic acids may be used for recombinant
CC production of BEBR, gene therapy, antisense therapy, as probes for
CC diagnostic assays and for functional studies. BEBR may be used to raise
CC antibodies and for identification of BEBR modulators. BEBR antagonists
CC may be used to increase bone mineral content for the treatment of
CC disorders such as osteopenia, osteoporosis, fractures and other
CC disorders associated with low mineral content.
XX
SQ Sequence 2301 BP; 568 A; 635 C; 614 G; 484 T; 0 other;
Query Match 78.9%; Score 501.6; DB 21; Length 2301;
Best Local Similarity 87.5%; Fred. No. 9.2e-96;
Matches 562; Conservative 0; Mismatches 74; Indels 6; Gaps 1;
Qy 1 ATGCAAGCCCTCACTAGCCCGGCTCATCTGCTTCTGTCAGAGCTTCTGCT 60
Db 48 ATGCAAGCCCTCACTAGCCCGGCTCATCTGCTTCTGTCAGAGCTTCTGCT 107
Qy 61 GTGAGAGGCGCAAGGCTGCGCAAGCTTCAAGAAATGATCCACAGAGTCAATCCAGAGCTT 120
Db 108 GTGAGAGGCGCAAGGCTGCGCAAGCTTCAAGAAATGATCCACAGAGTCAATCCAGAGCTT 167
Qy 121 GGAGAGTACCCCGAGGCTCTCC-----TGAGAACACCAAGCATGAAACCGGCGGAG 174
Db 168 GGAGAGTACCCCGAGGCTCTCCCGAGGCTCTCCCGAGGCTCTCCCGAGGCTCTCCCGAGG 227
Qy 175 AATGAGAGGAGACTTCCCAACCATCTTATGACGCGCAAGATGTGTCCAGTACAGCTGC 234
Db 228 AATGAGAGGAGACTTCCCAACCATCTTATGACGCGCGCGCGCTTGTGACCAAGAGTGTCCAGTACAGCTGC 287
Qy 235 CGCGAGTGCATTAACCGGCTTCTGACAGAGCGGCCATGCGGACGCGCAAGCGGTC 294
Db 288 CGCGAGTGCATTAACCGGCTTCTGACAGAGCGGCCATGCGGCGCGCGCAAGCGGTC 347
Qy 295 ACCGAGTTGTGTCTCCGCGCAAGTGGCGGCGCGCGGCGGCTGCTGCGCAAGTGTGCT 354
Db 348 ACCGAGTTGTGTCTCCGCGCAAGTGGCGGCGCGCGGCGGCTGCTGCGCAAGTGTGCT 407
Qy 355 CGCGTGAAGTGTGCGCGCGCGCAAGGACGATTTCCGCTGATCCCGATGCTAACCG 414
Db 408 CGCGTGAAGTGTGCGCGCGCGCAAGGACGATTTCCGCTGATCCCGATGCTAACCG 467
Qy 415 GCGCAGCGGCTGCACTGCTGTGCGCGGCGCGCGCGCGCGCTGCTGCGCAAGTGTGCT 474
Db 468 GCGCAGCGGCTGCACTGCTGTGCGCGGCGCGCGCGCGCGCGCGCGCAAGTGTGCT 527
Qy 475 CTGCTGCGCTTCTGTCAGAGTGCAGAGCGGCTTCAACCGCTTCAACAGTGCAGTCAAG 534
Db 528 CTGCTGCGCTTCTGTCAGAGTGCAGAGCGGCTTCAACCGCTTCAACAGTGCAGTCAAG 587
Qy 535 GACTTTCGCGCGCGGAGACCGCGCGCGCGCGCAAGGCTGCGCAAGCGGCGCGCGCG 594
Db 588 GACTTTCGCGCGCGGAGACCGCGCGCGCGCGCGCAAGGCTGCGCAAGCGGCGCGCGCG 647
Qy 595 GGAGCCAAAGCCAAACGAGCGGAGCTGAGAGAGCGCTTCTTAC 636
Db 648 GGAGCCAAAGCCAAACGAGCGGAGCTGAGAGAGCGCTTCTTAC 689

RESULT 9
AAA29056
ID AAA29056 standard; cDNA; 2301 BP.

```


XX AC AAA29056;
 XX 12-SEP-2000 (first entry)
 DE Human TGF-beta binding protein (BEER) variant V101 cDNA.
 XX osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KW BEER, variant; V101; gene therapy; antisense therapy; fracture;
 XX chromosome 17q12-21; bone mineralization; ss.
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FT CDS 48..689
 FT /tag= a
 FT /label= BEER variant V101
 FT /product= TGF-beta_binding_protein
 XX MO200032773-A1.
 XX 08-JUN-2000.
 XX 24-NOV-1999; 99WO-US27990.
 XX 27-NOV-1998; 98US-0110283.
 XX (DARW-) DARWIN DISCOVERY LTD.
 XX Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepker BW,
 XX Van Ness J, Winkler DG;
 XX WPI; 2000-412321/35.
 DR P-RSDB; AAY96430.
 PT Nucleic acids (I) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures
 XX
 PS Claim 1; Page 118-119; 162pp; English.
 XX
 CC This cDNA encodes a variant human transforming growth factor-beta
 CC (TGF-beta) binding protein designated BEER V101. The encoded protein
 CC comprises a substitution of isoleucine for the wild-type valine at
 CC residue 10. The cDNA and protein may be used for prevention, treatment,
 CC and diagnosis of diseases associated with inappropriate BEER expression.
 CC For example, they may be used to treat disorders associated with
 CC decreased TGF-beta BP expression. The cDNA or vectors may be administered
 CC to treat diseases by rectifying mutations or deletions in a patient's
 CC genome that affect the activity of BEER by expressing inactive proteins
 CC or to supplement the patient's own production of BEER polypeptides. The
 CC nucleic acids may be used for recombinant production of BEER. Gene
 CC therapy, antisense therapy, as probes for diagnostic assays and for
 CC functional studies. BEER may be used to raise antibodies and for
 CC identification of BEER modulators. BEER antagonists may be used to
 CC increase bone mineral content for the treatment of disorders such as
 CC osteopenia, osteoporosis, fractures and other disorders associated with
 CC low mineral content.
 CC
 XX Sequence 2301 BP; 569 A; 634 C; 614 G; 484 T; 0 other;
 SQ
 Query Match 78.9%; Score 501.6; DB 21; Length 2301;
 Best Local Similarity 87.5%; Pred. No. 9.2e-96;
 Matches 562; Conservative 0; Mismatches 74; Indels 6; Gaps 1;
 QY 1 ATGAGAGCCCTCACTAGCCCGCTGCTCATCTGCTACTTGTGACGCTGCTTCTGTGCT 60
 DB 48 ATGAGAGCTCCCACTGCGCTGTGTCTCATCTGCTGTGTAACAACAAGCTTCCGTGA 107
 QY 61 GTGAGAGGCCAGAGGGGTGACCACTTCAAGAGATGATCCACAGAGGTCAATCCAGAGGCTT 120
 DB 108 GTGAGAGGCCAGAGGGGTGACGCTTCAAGATATATCCACGAAATATATCCGAGAGCTC 167

QY 121 GGAGAGTACCCCGAGCCTCTCTCC-----TGAGAACACCAACCAATGAACCGGCGGAG 174
 DB 168 GGAGAGTACCCCGAGCCTCTCTCCACCGAGAGCTTGAAGAACAAAGACATGAACCGGCGGAG 227
 QY 175 AATGAGGCGAGACTCTCCACCATCTCTATGACGCGCAAAATGTGTCCAGTACAGTTC 234
 DB 228 AACGAGGCGCGCTCTCCACCAACCCCTTTGAGCCAAAGAGTGTCCAGTACAGTTC 287
 QY 235 CGCGAGCTGACTACACCGGCTCTCTGACAGAGGAGCCATATGCGGACGCGCAAGCGGCTC 294
 DB 288 CGGAGCTGACTACACCGGCTCTCTGACAGAGGAGCCATATGCGGACGCGCAAGCGGCTC 347
 QY 295 ACCGAGTTGATGTGCTCTCGGCAAGTGGCGGCGCGGCGGCTCTCTGCAACCGCATCGG 354
 DB 348 ACCGAGCTGATGTGCTCTCGGCAAGTGGCGGCGGCGGCGGCTCTCTGCAACCGCATCGG 407
 QY 355 CGCGTGAAGTGTGCGCGCGGCAAGGACCGGATTTCCGCTGATCCCGGATCCGTACCGC 414
 DB 408 CGCGGCAAGTGTGCGCGCGGCAAGGAGGAGCCGACTTCCGCTGATCCCGGATCCGTACCGC 467
 QY 415 GCGGAGCGGAGTGAAGCTGCTGCGCGCGGAGGCGCGGCGGCGGCTCTCTGCAAGGAGTGC 474
 DB 468 GCGGAGCGGAGTGAAGCTGCTGCGCGGAGGCGCGGCGGCGGCTCTCTGCAAGGAGTGC 527
 QY 475 CTGATGAGCTCTGTCGAAGTGAAGCGGCTCAACCGCTTCAACAACGATCGAGCTCAAG 534
 DB 528 CTGATGAGCTCTGTCGAAGTGAAGCGGCTCAACCGCTTCAACAACGATCGAGCTCAAG 587
 QY 535 GACTTTCGGGCGGAGACCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 594
 DB 588 GACTTTCGGGCGGAGACCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 647
 QY 595 GGAGCCAAAGCCAAACAGAGCGGAGCTGAGAACGCTTACTAG 636
 DB 648 AGCGCCAAAGCCAAACAGAGCGGAGCTGAGAACGCTTACTAG 689
 RESULT 10
 ID AAA91023
 XX AAA91023 standard; DNA; 2329 BP.
 XX
 AC AAA91023;
 XX
 DT 05-APR-2001 (first entry)
 XX
 DE Human secreted protein PRO7476 coding sequence.
 KW Secreted protein; human; PRO protein; neoplastic cell growth; tumour;
 KW proliferation; leukaemia; lymphoid malignancy; inflammatory disorder;
 KW angiogenic disorder; immunologic disorder; PRO7476; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 62..703
 FT /tag= a
 FT /product= PRO7476
 XX
 PN MO200075317-A2.
 XX
 PD 14-DEC-2000.
 XX
 PF 15-MAY-2000; 2000WO-US13358.
 XX
 PR 09-JUN-1999; 99US-0138385.
 PR 20-JUN-1999; 99US-0144790.
 PR 03-AUG-1999; 99US-0146843.
 PR 10-AUG-1999; 99US-0148188.
 PR 17-AUG-1999; 99US-0149320.
 PR 17-AUG-1999; 99US-0149327.
 PR 17-AUG-1999; 99US-0149396.
 PR 20-AUG-1999; 99US-0150114.
 PR 31-AUG-1999; 99US-0151700.

invention also comprises a method for producing the proteins of the invention by recombinant means and antibodies specific for the protein of the invention. The antibody may be used for detecting the PRO proteins of the invention and may be used to modify their activity. CC polynucleotides may be used as hybridisation probes for a cDNA library to isolate the full-length PRO cDNA or to isolate other cDNAs, to construct hybridisation probes for mapping the gene which encodes that PRO and for genetic analysis of individuals with genetic disorders, in assays to identify other proteins or molecules involved in binding reaction, to generate transgenic animals or knock-out animals which in turn are useful in the development and screening of therapeutically useful reagents, for chromosome identification, and tissue typing. The PRO polypeptides are useful in gene therapy, and as molecular weight markers for protein electrophoresis purposes. The sequences may also be used to detect overexpression on PRO polypeptides in cancerous tumours and for screening for differentially expressed genes using microarray technology. The present sequence represents a cDNA encoding a human PRO protein of the invention.

Sequence 2329 BP; 587 A; 645 C; 612 G; 485 T; 0 other;

Query Match 78.9%; Score 501.6; DB 24; Length 2329;
Best Local Similarity 87.5%; Pred. No. 9.2e-95;
Matches 562; Conservative 0; Mismatches 74; Indels 6; Gaps 1;

1 ATGCAGCCCTCACTAGCCCGGCTGCTCATCTGCTTGTGACGCTGCTTCTGTCT 60
62 ATGCAGCTCCCACTAGCCCTGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 121
61 GTGAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
122 GTGAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 181
121 GGAGAGTACCCCGAGGCTCTCTCC-----TGAGAACACACGAGCATGAGCCGAGGAG 174
182 GGAGAGTACCCCGAGGCTCTCTCC-----TGAGAACACACGAGCATGAGCCGAGGAG 241
175 AATGAGGAGCAGACTCTCCCACTGCTTATGACGCGCAAGATGTGTCGAGTACAGCTGC 234
242 AACGAGGAGGAGGCTCTCCCACTGCTTATGACGCGCAAGATGTGTCGAGTACAGCTGC 301
235 CGGAGGCTGCACTACACCGGCTTCTGACAGAGCGCCGAGGAGGAGGAGGAGGAGGAG 284
302 CGGAGGCTGCACTACACCGGCTTCTGACAGAGCGCCGAGGAGGAGGAGGAGGAGGAG 361
295 ACCGAGTGTGTGCTCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 354
362 ACCGAGTGTGTGCTCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 421
355 CGGAGTGTGTGCTCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 414
422 CGGAGTGTGTGCTCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 481
415 GCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 474
482 GCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 541
475 CTGAGTGTGTGCTCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 534
542 CTGAGTGTGTGCTCTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 601
535 GACTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 594
602 GACTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 661
595 GAGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 636
662 AGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 703

RESULT 12
AAA29061
ID AAA29061 standard; cDNA; 2301 BP.

AC AAA29061;
XX
DT 12-SEP-2000 (first entry)
XX
DE Mutant human TGF-beta binding protein (BEER) cDNA.
XX
KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
KW BEER; chromosome 17q12-21; gene therapy; antisense therapy; fracture;
KW bone mineralization; mutant; sclerosteosis; ss.
XX
OS Homo sapiens.
XX
FH Key
FH CDS Location/Qualifiers
FT 48..119
FT /*tag= a
FT /note= "mutant BEER created by sclerosteosis
FT nonsense mutation"

W0200032773-A1.

08-JUN-2000.

24-NOV-1999; 99MO-US27990.

27-NOV-1998; 98US-0110283.

(DARW-) DARWIN DISCOVERY LTD.

Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW,
Van Ness J, Winkler DG;

WPI; 2000-412321/35.
P-PSDB; AAY96435.

Nucleic acid (1) encoding a transforming growth factor beta binding protein, useful for identifying agents for treating osteopenia, osteoporosis and fractures
Claim 27; Page 117-118; 162pp; English.

This cDNA encodes a mutant human transforming growth factor-beta (TGF-beta) binding protein (BEER) produced as a result of a nonsense mutation in the BEER coding sequence (C to T mutation at position 117). This mutation has been linked to sclerosteosis in humans by linkage analysis of an Afrikaaner population in South Africa. The hBEER gene has been localized to the chromosome 17q12-21. The cDNA and protein may be used for prevention, treatment and diagnosis of diseases associated with inappropriate BEER expression. For example, they may be used to treat disorders associated with decreased TGF-beta BP expression. The cDNA or vectors may be administered to treat diseases by rectifying mutations or deletions in a patient's genome that affect the activity of BEER by expressing inactive proteins or to supplement the patient's own production of BEER polypeptides. The nucleic acids may be used for recombinant production of BEER, gene therapy, antisense therapy, as probes for diagnostic assays and for functional studies. BEER may be used to raise antibodies and for identification of BEER modulators. BEER antagonists may be used to increase bone mineral content for the treatment of disorders such as osteopenia, osteoporosis, fractures and other disorders associated with low mineral content.

Sequence 2301 BP; 568 A; 634 C; 614 G; 485 T; 0 other;

Query Match 78.6%; Score 500; DB 21; Length 2301;
Best Local Similarity 87.4%; Pred. No. 2e-95;
Matches 561; Conservative 0; Mismatches 75; Indels 6; Gaps 1;

1 ATGCAGCCCTCACTAGCCCGGCTGCTCATCTGCTTGTGACGCTGCTTGTGCT 60
48 ATGCAGCTCCCACTAGCCCTGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 107
61 GTGAGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120

Db 108 GTGAGAGGCTAGAGGAGTGGCAGAGCGCTTCAGAGATGATGCCAGAAATCATCCCGAGCTC 167
 Qy 121 GGAGAGTACCCCGAGAGCTCTCTCC-----TGAGAACCAACGAGCATGAAACGGGGGAG 174
 Db 168 GGAGAGTACCCCGAGAGCTCTCTCCAGAGAGCTGAGAGAACAAAGACATGAACCGGGGAG 227
 Qy 175 AATGAGAGCAGACCTTCCCAACATCCCTATGAGCGCCAAAGATGTGTCCAGTACAGCTGC 234
 Db 228 AACGAGAGGCGGCTCCCAACCAACCTTTGAGACCAAAACGATGTCCAGTACAGCTGC 287
 Qy 235 CGGAGAGTGCATACACCCGCTTCTGACAGAGGCCCATGCGGCAAGCCGCTC 294
 Db 288 CGGAGAGTGCATACACCCGCTTCTGACAGAGGCCCATGCGGCAAGCCGCTC 347
 Qy 295 ACCGAGTGTGTGTCTCCGCGCAGTGGCGCCCGCGCGGTGTGTCGCCAACGCGCATCGAG 354
 Db 348 ACCGAGTGTGTGTCTCCGCGCAGTGGCGCCCGCGCGGTGTGTCGCCAACGCGCATCGAG 407
 Qy 355 CGCGTGAAGTGTGTGTCTCCGCGCAGTGGCGCCCGCGCGGTGTGTCGCCAACGCGCATCGAG 414
 Db 408 CGCGGCAAGTGTGTGTCTCCGCGCAGTGGCGCCCGCGCGGTGTGTCGCCAACGCGCATCGAG 467
 Qy 415 GCGCAGAGGAGTGCAGTGTGTGTCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 474
 Db 468 GCGCAGAGGAGTGCAGTGTGTGTCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 527
 Qy 475 CTGAGTGGCTCTGTGCAAGTGCAGAGCGCTTCCAGACGAGTGGAGCTCAAG 534
 Db 528 CTGAGTGGCTCTGTGCAAGTGCAGAGCGCTTCCAGACGAGTGGAGCTCAAG 587
 Qy 535 GACTTGGGCGCGGAGACCG 594
 Db 588 GACTTGGGCGCGGAGACCG 647
 Qy 595 GAGGCCAAGCCAAACGAGCGCGAGCTGAGAGAGCTTACTAG 636
 Db 648 AGCGCAAAAGCCAAACGAGCGCGAGCTGAGAGAGCTTACTAG 689
 RESULT 13
 AAA29062
 ID AAA29062 standard; cDNA, 2301 BP.
 AC AAA29062;
 XX
 DT 12-SEP-2000 (first entry)
 XX
 DE Human TGF-beta binding protein (BBER) variant P38R cDNA.
 XX
 KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KM BBER; variant; P38R; gene therapy; antisense therapy; fracture;
 KM chromosome 17q12-21; bone mineralization; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 48..689
 FT /tag= a
 FT /label= BBER variant_P38R
 FT /product= TGF-beta_binding_protein
 XX
 PN WO200032773-A1.
 PD 08-JUN-2000.
 XX
 PF 24-NOV-1999; 99WO-US27990.
 XX
 PR 27-NOV-1998; 98US-0110283.
 XX
 PA (DARW-) DARWIN DISCOVERY LTD.
 XX
 PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW,
 Van Ness J, Winkler DG;

XX WPI: 2000-412321/35.
 DR P-PSDB; AAY96436#.
 XX
 PT Nucleic acid (1) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures
 XX
 PS Claim 27; Page 120-121; 162pp; English.
 XX
 CC This cDNA encodes a variant human transforming growth factor-beta
 CC (TGF-beta) binding protein designated BBER P38R. The encoded protein
 CC comprises a substitution of arginine for the wild-type proline at
 CC residue 38. The cDNA and protein may be used for prevention, treatment,
 CC and diagnosis of diseases associated with inappropriate BBER expression.
 CC For example, they may be used to treat disorders associated with
 CC decreased TGF-beta BP expression. The cDNA or vectors may be administered
 CC to treat diseases by rectifying mutations or deletions in a patient's
 CC genome that affect the activity of BBER by expressing inactive proteins
 CC or to supplement the patient's own production of BBER polypeptides. The
 CC nucleic acids may be used for recombinant production of BBER, gene
 CC therapy, antisense therapy, as probes for diagnostic assays and for
 CC functional studies. BBER may be used to raise antibodies and for
 CC identification of BBER modulators. BBER antagonists may be used to
 CC increase bone mineral content for the treatment of disorders such as
 CC osteopenia, osteoporosis, fractures and other disorders associated with
 CC low mineral content.
 XX
 SQ Sequence 2301 BP; 568 A; 634 C; 615 G; 484 T; 0 other;
 Query Match 78.6%; Score 500; DB 21; Length 2301;
 Best Local Similarity 87.4%; Pred. No. 2e-95;
 Matches 561; Conservative 0; Mismatches 75; Indels 6; Gaps 1;
 Qy 1 ATGCAAGCCCTCACTAGCCCGCTGCTCTATGCTTGTGACGCTGCTTGTGCT 60
 Db 48 ATGCAAGCTCCCACTGAGCCCTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 107
 Qy 61 GTGAGAGGCGCAGGGGTGTGCAAGCCCTTCAAGATATGCAAGAGGTATCCAGGGCTT 120
 Db 108 GTGAGAGGCGCAGGGGTGTGCAAGCCCTTCAAGATATGCAAGAGGTATCCAGGGCTT 167
 Qy 121 GGAGAGTACCCCGAGAGCTCTCTCC-----TGAGAACCAACGAGCATGAAACGGGGGAG 174
 Db 168 GGAGAGTACCCCGAGAGCTCTCTCCAGAGAGCTGAGAGAACAAAGACATGAACCGGGGAG 227
 Qy 175 AATGAGAGCAGACCTTCCCAACATCCCTATGAGCGCCAAAGATGTGTCCAGTACAGCTGC 234
 Db 228 AACGAGAGGCGGCTCCCAACCAACCTTTGAGACCAAAACGATGTCCAGTACAGCTGC 287
 Qy 235 CGGAGAGTGCATACACCCGCTTCTGACAGAGGCCCATGCGGCAAGCCGCTC 294
 Db 288 CGGAGAGTGCATACACCCGCTTCTGACAGAGGCCCATGCGGCAAGCCGCTC 347
 Qy 295 ACCGAGTGTGTGTCTCCGCGCAGTGGCGCCCGCGCGGTGTGTCGCCAACGCGCATCGAG 354
 Db 348 ACCGAGTGTGTGTCTCCGCGCAGTGGCGCCCGCGCGGTGTGTCGCCAACGCGCATCGAG 407
 Qy 355 CGCGTGAAGTGTGTGTCTCCGCGCAGTGGCGCCCGCGCGGTGTGTCGCCAACGCGCATCGAG 414
 Db 408 CGCGGCAAGTGTGTGTCTCCGCGCAGTGGCGCCCGCGCGGTGTGTCGCCAACGCGCATCGAG 467
 Qy 415 GCGCAGAGGAGTGCAGTGTGTGTCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 474
 Db 468 GCGCAGAGGAGTGCAGTGTGTGTCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 527
 Qy 475 CTGAGTGGCTCTGTGCAAGTGCAGAGCGCTTCCAGACGAGTGGAGCTCAAG 534
 Db 528 CTGAGTGGCTCTGTGCAAGTGCAGAGCGCTTCCAGACGAGTGGAGCTCAAG 587
 Qy 535 GACTTGGGCGCGGAGACCG 594
 Db 588 GACTTGGGCGCGGAGACCG 647

QY 595 GGAGCCAAAGCAACGAGCTTGAGAGCCCTACTAG 636
 DB 648 AGGCGCAAGCAACGAGCTTGAGAGCCCTACTAG 689

RESULT 14

AAA29063
 ID AAA29063 standard; DNA; 35828 BP.

AC AAA29063;

DT 12-SEP-2000 (first entry)

DE Murine TGF-beta binding protein (BBER) genomic MluI-AvII DNA fragment.

KW Osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KM BBER; gene therapy; antisense therapy; fracture; bone mineralization;
 XX transgenic mouse; over-expression; ds.

OS Mus musculus.

XX Key Location/Qualifiers

FT CDS 18317..21464

FT /*tag= a

FT 18317..18530

FT /*tag= b

FT intron 18531..21022

FT /*tag= c

FT exon 21023..21464

FT /*tag= d

FT /partial

XX WO200032773-A1.

XX 08-JUN-2000.

XX 24-NOV-1999; 99WO-US27990.

XX 27-NOV-1998; 98US-0110283.

XX (DARW-) DARWIN DISCOVERY LTD.

XX PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;

XX XI Van Ness J, Winkler DG;

XX DR WPI; 2000-412321/35.

XX P-PSDB; AAY96432.

XX PT Nucleic acids (1) encoding a transforming growth factor beta binding

XX PT protein, useful for identifying agents for treating osteopenia,

XX PT osteoporosis and fractures

XX PS Example 9; Page 74-87; 162pp; English.

XX This 35 kb MluI-AvII genomic DNA restriction fragment includes the

XX entire murine transforming growth factor-beta (TGF-beta) binding protein

XX (BBER) gene. This DNA was used to generate a transgenic mouse

XX over-expressing the BBER gene. The gene and protein may be used for

XX prevention, treatment and diagnosis of diseases associated with

XX inappropriate BBER expression. For example, they may be used to treat

XX disorders associated with decreased TGF-beta BP expression. The cDNA or

XX vectors may be administered to treat diseases by rectifying mutations or

XX deletions in a patient's genome that affect the activity of BBER by

XX expressing inactive proteins or to supplement the activity of BBER by

XX production of BBER polypeptides. The nucleic acids may be used for recombinant

XX production of BBER, gene therapy, antisense therapy, as probes for

XX diagnostic assays and for functional studies. BBER may be used to raise

XX antibodies and for identification of BBER modulators. BBER antagonists

XX may be used to increase bone mineral content for the treatment of

XX disorders such as osteopenia, osteoporosis, fractures and other disorders

XX associated with low mineral content.

XX SQ Sequence 35828 BP; 8973 A; 8817 C; 9064 G; 8973 T; 1 other;

Query Match 67.1%; Score 426.6; DB 21; Length 35828;

Best Local Similarity 95.8%; Pred. No. 4, 5e-80;

Matches 438; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 180 AGGCGAGACCTCCCAACATCCCTATGACGCCCAAGATGTGTCCGATGACGCTCCCGCA 239

DB 20988 AGGCTGCAGCCCTTCACGATCCTCTCTCCGCAAGATGTGTCCGATGACGCTCCCGCA 21047

QY 240 GCTGCATACACCCGCTTCCTGACAGAGGCCATGCGCGCAAGCCGATCCTACCGA 299

DB 21048 GCTGCATACACCCGCTTCCTGACAGAGGCCATGCGCGCAAGCCGATCCTACCGA 21107

QY 300 GTTGTGTGTCTCCGCGCAGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 359

DB 21108 GTTGTGTGTCTCCGCGCAGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 21167

QY 360 GAAGTGTGTGCG 419

DB 21168 GAAGTGTGTGCG 21227

QY 420 GCGGCTGTCAGCTGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 479

DB 21228 GCGGCTGTCAGCTGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 21287

QY 480 GGGCTGTGTGCAAGTGCAGAGCGCGCTCACCGCGTTCACACAGTGCAGAGTCAAGGACT 539

DB 21288 GGGCTGTGTGCAAGTGCAGAGCGCGCTCACCGCGTTCACACAGTGCAGAGTCAAGGACT 21347

QY 540 CGGCGCGGAGACCG 599

DB 21348 CGGCGCGGAGACCG 21407

QY 600 CAAAGCCAAACGAGCGGAGCTTGAGAGCCCTACTAG 636

DB 21408 CAAAGCCAAACGAGCGGAGCTTGAGAGCCCTACTAG 21444

RESULT 15

AAA29060
 ID AAA29060 standard; cDNA; 532 BP.

AC AAA29060;

DT 12-SEP-2000 (first entry)

DE Bovine TGF-beta binding protein (BBER) cDNA.

KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KM BBER; gene therapy; antisense therapy; fracture; bone mineralization; ss.

OS Bos taurus.

XX Key Location/Qualifiers

FT CDS 1..530

FT /*tag= a

FT /partial

FT /product= TGF-beta_binding_protein

XX WO200032773-A1.

XX 08-JUN-2000.

XX 24-NOV-1999; 99WO-US27990.

XX 27-NOV-1998; 98US-0110283.

XX (DARW-) DARWIN DISCOVERY LTD.

XX PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;

XX XI Van Ness J, Winkler DG;

XX
DR WPI: 2000-412321/35.
P-PSDB; AA196434.

XX
PT Nucleic acids (1) encoding a transforming growth factor beta binding
PT protein, useful for identifying agents for treating osteopenia,
PT osteoporosis and fractures

PS
PS Claim 1; Page 126; 162pp; English.

XX
CC This cDNA encodes a bovine transforming growth factor-beta (TGF-beta)
CC binding protein designated bBER. The cDNA and protein may be used for
CC prevention, treatment and diagnosis of diseases associated with
CC inappropriate bBER expression. For example, they may be used to treat
CC disorders associated with decreased TGF-beta BP expression. The cDNA or
CC vectors may be administered to treat diseases by rectifying mutations or
CC deletions in a patient's genome that affect the activity of bBER by
CC expressing inactive proteins or to supplement the patient's own production
CC of bBER polypeptides. The nucleic acids may be used for recombinant
CC production of bBER, gene therapy, antisense therapy, as probes for
CC diagnostic assays and for functional studies. bBER may be used to raise
CC antibodies and for identification of bBER modulators. bBER antagonists
CC may be used to increase bone mineral content for the treatment of
CC disorders such as osteopenia, osteoporosis, fractures and other disorders
CC associated with low mineral content.

XX
SQ Sequence 532 BP; 95 A; 199 C; 176 G; 62 T; 0 other;

Query Match 65.4%; Score 415.8; DB 21; Length 532;

Best Local Similarity 87.8%; Pred. No. 6,7e-78;

Matches 466; Conservative 0; Mismatches 62; Indels 3; Gaps 1;

QY 90 GAATGATGCCACAGAGTCAATCCAGGAGCTTGGAGAGTACCCGAGCCTCTCTCTGA--- 146
DB 2 GAATGATGCCACAGAAATCAATCCCGAGCTGGGCGAGTACCCGAGCCTCTGCGAGAGCT 61
QY 147 GAACAAACGAGCACTGAACCGGGGCGAGAAATGAGGAGACCTCCCAACATCCCTTGA 206
DB 62 GAACAAACGAGCACTGAACCGGGGCGAGAAATGAGGAGACCTCCCAACATCCCTTGA 121
QY 207 CGCCAAAGATGTCTCGAGTACAGTCCGCGAGCTGCACTACACCGCTTCTCTGACAGA 266
DB 122 GACCAAGAGCGCTCCGAGTACAGTCCGCGAGCTGCACTACACCGCTTCTCTGACAGA 181
QY 267 CGGCGCATGCCGAGCGCCCAAGCCGATCAGGAGTGTGCTCCGACCAAGTCCGCC 326
DB 182 TGGGCGGTGCGGAGCGCCCAAGCCGATCAGGAGTGTGCTCCGACCAAGTCCGCC 241
QY 327 CGGCGCGTGTGCGGAGCGCCCAAGCCGATCAGGAGTGTGCTCCGACCAAGTCCGCC 386
DB 242 GGGCGGCTGTGCGGAGCGCCCAAGCCGATCAGGAGTGTGCTCCGACCAAGTCCGCC 301
QY 387 TTTCCGCTGATCCCGGATGCTACCGCGGAGCGGAGTGTGCTGCTGCTGCTGCTGCTGCT 446
DB 302 CTTCGCTGATCCCGGATGCTACCGCGGAGCGGAGTGTGCTGCTGCTGCTGCTGCTGCT 361
QY 447 CGGCGCGTGTGCGGAGCGCCCAAGCCGATCAGGAGTGTGCTGCTGCTGCTGCTGCTGCT 506
DB 362 CGGCGCGTGTGCGGAGCGCCCAAGCCGATCAGGAGTGTGCTGCTGCTGCTGCTGCTGCT 421
QY 507 CCGCTTCCAGCAACAGTGTGAGTCAAGAGCTTGGGCGGAGACCGGCGGCGAGAA 566
DB 422 TCGCTTCCAGCAACAGTGTGAGTCAAGAGCTTGGGCGGAGACCGGCGGCGAGAA 481
QY 567 GGGTCCAGACCGGCGGCGGCGGAGACCGGCGGAGACCGGCGGAGACCGGCGGAGACCGG 617
DB 482 GGGCGGAGAGTGTGCGGCGGCGGAGACCGGCGGAGACCGGCGGAGACCGGCGGAGACCG 532

Search completed: March 28, 2003, 23:23:25
Job time : 215.234 secs

•
•
•
•

0

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2003, 00:44:05 ; Search time 2065.75 Seconds

(without alignments)
8960.144 Million cell updates/sec

Title: US-09-867-274-3

Perfect score: 636
Sequence: 1 atcgagccctcactagcccc.....agctgagagcgcctactag 636Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

```
1: gb_da:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_stc:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_stc:*
28: em_un:*
29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pln:*
35: em_hcg_rod:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hcgo_hum:*
40: em_hcgo_mus:*
41: em_hcgo_other:*
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	636	100.0	636	6 AX323455	AX323455 Sequence
2	585	92.0	638	10 AF326740	AF326740 Mus muscu
3	424	66.7	81806	10 AF326737	AF326737 Mus muscu
4	424	66.7	110000	2 AC068782_2	AC068782_2
5	424	66.7	205277	2 AC012296	Continuation (3 of
6	424	66.7	208135	10 AC068807	AC068807 Mus muscu
7	373	58.6	198508	10 AL591145	AL591145 Mouse DNA
8	214	33.6	51575	2 AC023810	AC023810 Mus muscu
9	104	16.4	674	10 AF326741	AF326741 Rattus no
10	104	16.4	101804	2 AC098160	AC098160 Rattus no
11	104	16.4	104898	2 AC121721	AC121721 Rattus no
12	69	10.8	759	6 AX323453	AX323453 Sequence
13	69	10.8	2271	6 AX342535	AX342535 Sequence
14	69	10.8	2296	9 AF331844	AF331844 Homo sapi
15	69	10.8	2323	9 AF326739	AF326739 Homo sapi
16	69	10.8	2329	6 AX056687	AX056687 Sequence
17	69	10.8	7099	6 AX342537	AX342537 Sequence
18	69	10.8	21501	9 AF326736	AF326736 Homo sapi
19	69	10.8	80117	9 AC055813	AC055813 Homo sapi
20	69	10.8	94752	9 AC003098	AC003098 Homo sapi
21	69	10.8	177744	2 AC073954	AC073954 Homo sapi
22	68	10.7	642	9 AF326742	AF326742 Homo sapi
23	32	5.0	532	4 AF326738	AF326738 Bos tauru
24	26	4.1	41	6 AX323472	AX323472 Sequence
25	25	3.9	25	6 AX323466	AX323466 Sequence
26	25	3.9	45	6 AX323471	AX323471 Sequence
27	24	3.8	24	6 AX323464	AX323464 Sequence
28	21	3.3	21	6 AX323467	AX323467 Sequence
29	21	3.3	21	6 AX323469	AX323469 Sequence
30	21	3.3	27	6 AX056700	AX056700 Sequence
31	21	3.3	29	6 AX323461	AX323461 Sequence
32	21	3.3	3552	9 AB002058	AB002058 Homo sapi
33	21	3.3	28984	9 AB002059	AB002059 Homo sapi
34	21	3.3	147086	9 AC002472	AC002472 Homo sapi
35	21	3.3	152172	8 AC068924	AC068924 Oryza sat
36	21	3.3	162470	9 AC007664	AC007664 Homo sapi
37	20	3.1	24	6 AX323462	AX323462 Sequence
38	20	3.1	10460	1 AE011791	AE011791 Xanthomon
39	20	3.1	14108	1 AE004989	AE004989 Halobacte
40	20	3.1	45653	2 AC105239	AC105239 Homo sapi
41	20	3.1	69337	2 AC101261	AC101261 Mus muscu
42	20	3.1	97216	2 AC121752	AC121752 Rattus no
43	20	3.1	97684	9 HS309122	HS309122 Human DNA s
44	20	3.1	136081	2 AC128778	AC128778 Rattus no
45	20	3.1	139998	2 AC111270	AC111270 Rattus no

ALIGNMENTS

RESULT 1
AX323455
LOCUS AX323455
DEFINITION Sequence 3 from Patent WO01923308.
ACCESSION AX323455
VERSION AX323455.1 GI:18094217
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Paszty, C.J. and Gao, Y.
Cysteine-knot polypeptides: cloaked-2 molecules and uses thereof
Patent: WO 01923308-A 3 06-DEC-2001;

Amgen, Inc. (US)
Location/Qualifiers
1. .636
/organism="Mus musculus"
/db_xref="taxon:10090"

BASE COUNT 114 a 224 c 207 g 91 t

ORIGIN

Query Match 100.0%; Score 636; DB 6; Length 636;
Best Local Similarity 100.0%; Pred. No. 8.3e-305;
Matches 636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGAGAGCCCTCACTAGAGCCGCTGCTCTATCTGCTTATTTTGACAGCTGCTCTGCTTGTGCT 60
1 ATGAGAGCCCTCACTAGAGCCGCTGCTCTATCTGCTTATTTTGACAGCTGCTCTGCTTGTGCT 60

61 GTGAGAGGCGAGGGGTGGCAAGCTTTCAGGAATGATCCACAGAGGTCATCCAGGGCTT 120
61 GTGAGAGGCGAGGGGTGGCAAGCTTTCAGGAATGATCCACAGAGGTCATCCAGGGCTT 120

121 GGAGAGTACCCCGAGCCTCTCTGAGAACACCAACATGAAACCGGGCGAGAAATGA 180
121 GGAGAGTACCCCGAGCCTCTCTGAGAACACCAACATGAAACCGGGCGAGAAATGA 180

181 GGAGAGCTCCCAACCTCCATAGAGCCAAAGATGTGTCAGATGACAGCTCCCGGAG 240
181 GGAGAGCTCCCAACCTCCATAGAGCCAAAGATGTGTCAGATGACAGCTCCCGGAG 240

241 CTGACACTACACCCGCTCTCTGACAGAGCCCATGCGCAGCCGCAAGCGGTCAACCGAG 300
241 CTGACACTACACCCGCTCTCTGACAGAGCCCATGCGCAGCCGCAAGCGGTCAACCGAG 300

301 TTGAGTGTCTCCGCGCAAGTGGGCGCGCGCGCTGCTGCCCAACCGCATGCGGCGGTG 360
301 TTGAGTGTCTCCGCGCAAGTGGGCGCGCGCGCTGCTGCCCAACCGCATGCGGCGGTG 360

361 AAGTGTGGGCGCGCGCAAGCGGATTTTCCGCTGATCCCGGATGCGTCAACCGCGGAG 420
361 AAGTGTGGGCGCGCGCAAGCGGATTTTCCGCTGATCCCGGATGCGTCAACCGCGGAG 420

421 CGGAGTGCAGCTGCTGTGCGCGCGGGGCGCGCGCGCTCGCGCAAGTGTGCTGTG 480
421 CGGAGTGCAGCTGCTGTGCGCGGGGCGCGCGCGCTCGCGCAAGTGTGCTGTG 480

481 GCTGTGTGCAAGTGCAGAGCGCTTCAACCAACAGTGTGCAAGTGTGCAAGTGTG 540
481 GCTGTGTGCAAGTGCAGAGCGCTTCAACCAACAGTGTGCAAGTGTGCAAGTGTG 540

541 GGGCGGAGACCGCGCGCGCGCAAGAGGTGCAAGCGCGCGCGCGCGCGCGCGGAGCC 600
541 GGGCGGAGACCGCGCGCGCGCAAGAGGTGCAAGCGCGCGCGCGCGCGCGCGGAGCC 600

601 AAAGCCAAACGAGCGGAGCTGGAAGACGCTTACTAG 636
601 AAAGCCAAACGAGCGGAGCTGGAAGACGCTTACTAG 636

RESULT 2
AF326740 638 bp mRNA linear ROD 28-FEB-2001
LOCUS Mus musculus sclerostin mRNA, complete cds.
DEFINITION AF326740
ACCESSION AF326740.1 GI:13161022
VERSION
KEYWORDS
SOURCE Mus musculus.
ORGANISM Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 638)
Brunkow M.E., Gardner J.C., Van Ness J., Paepel B.W., Kovacevich B.R., Prohl S., Skonier J.E., Zhao L., Sabo P.J., Fu Y.H., Altsch R.S., Gillet L., Colbert T., Tacconi P., Galas D., Hamerema H., Beighton P. and Mulligan J.T.

TITLE Bone dysplasia sclerosteosis results from loss of the SOST gene product, a novel cystine knot-containing protein
JOURNAL Am. J. Hum. Genet. 68 (3), 577-589 (2001)
MEDLINE 21090529
PUBMED 11179006
REFERENCE 2 (bases 1 to 638)
AUTHORS Brunkow M.E., Gardner J.C., Van Ness J., Paepel B.W., Kovacevich B.R., Prohl S., Skonier J.E., Zhao L., Sabo P.J., Fu Y.H., Altsch R.S., Gillet L., Colbert T., Tacconi P., Galas D., Hamerema H., Beighton P. and Mulligan J.T.
TITLE Direct Submission
JOURNAL Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631 220th St. SE, Bothell, WA 98021, USA

FEATURES
source
1. .636
/organism="Mus musculus"
/strain="129Sv/J"
/db_xref="taxon:10090"
/chromosome="11"

CDS
1. .636
/codon_start=1
/product="sclerostin"
/protein_id="AAK13455.1"
/db_xref="GI:13161023"
/translation="MPSLAPCLICLVHAAFCVAVGGQVQAFRNDATVETPLGEYPPENNQTMRAENGRPPHPYDAGKVSYSCEHLHYRPLTDGCRSAKVELVCSGCCPRLPLPNAIGHVKWRRPENGDFRITPRYARQVQLCPGGA PRSKRLVASCCKRLTFHNSLEKDFEPSTARQKRPKRPARKAKVAQALENAY"

BASE COUNT 114 a 224 c 209 g 91 t

ORIGIN

Query Match 92.0%; Score 585; DB 10; Length 638;
Best Local Similarity 99.8%; Pred. No. 1.8e-279; Indels 0; Gaps 0;
Matches 635; Conservative 0; Mismatches 1;

1 ATGAGAGCCCTCACTAGAGCCGCTGCTCTATCTGCTTATTTTGACAGCTGCTCTGCTTGTGCT 60
1 ATGAGAGCCCTCACTAGAGCCGCTGCTCTATCTGCTTATTTTGACAGCTGCTCTGCTTGTGCT 60

61 GTGAGAGGCGAGGGGTGGCAAGCTTTCAGGAATGATCCACAGAGTGTATCCAGGGCTT 120
61 GTGAGAGGCGAGGGGTGGCAAGCTTTCAGGAATGATCCACAGAGTGTATCCAGGGCTT 120

121 GGAGAGTACCCCGAGCCTCTCTGAGAACACCAAGCATGAAACCGGGCGAGAAATGA 180
121 GGAGAGTACCCCGAGCCTCTCTGAGAACACCAAGCATGAAACCGGGCGAGAAATGA 180

181 GGAGAGTACCCCGAGCCTCTCTGAGAACACCAAGCATGAAACCGGGCGAGAAATGA 180
181 GGAGAGTACCCCGAGCCTCTCTGAGAACACCAAGCATGAAACCGGGCGAGAAATGA 180

241 CTGACACTACACCCGCTCTCTGACAGAGCCCATGCGCAGCCGCAAGCGGTCAACCGAG 300
241 CTGACACTACACCCGCTCTCTGACAGAGCCCATGCGCAGCCGCAAGCGGTCAACCGAG 300

301 TTGAGTGTCTCCGCGCAAGTGGGCGCGCGCGCTGCTGCCCAACCGCATGCGGCGGTG 360
301 TTGAGTGTCTCCGCGCAAGTGGGCGCGCGCGCTGCTGCCCAACCGCATGCGGCGGTG 360

361 AAGTGTGGGCGCGCGCAAGCGGATTTTCCGCTGATCCCGGATGCGTCAACCGCGGAG 420
361 AAGTGTGGGCGCGCGCAAGCGGATTTTCCGCTGATCCCGGATGCGTCAACCGCGGAG 420

421 CGGAGTGCAGCTGCTGTGCGCGGGGCGCGCGCGCTCGCGCAAGTGTGCTGTG 480
421 CGGAGTGCAGCTGCTGTGCGCGGGGCGCGCGCGCTCGCGCAAGTGTGCTGTG 480

481 GCTGTGTGCAAGTGCAGAGCGCTTCAACCAACAGTGTGCAAGTGTGCAAGTGTG 540
481 GCTGTGTGCAAGTGCAGAGCGCTTCAACCAACAGTGTGCAAGTGTGCAAGTGTG 540

541 GGGCGGAGACCGCGCGCGCGCAAGAGGTGCAAGCGCGCGCGCGCGCGCGGAGCC 600
541 GGGCGGAGACCGCGCGCGCGCAAGAGGTGCAAGCGCGCGCGCGCGCGCGGAGCC 600

Db 541 GGGCCGAGACCGCGCGCGGCGAGAGGGTCGCAAGCCGCGCGCCGCCCGGAGACC 600
Qy 601 AAGCCCAACGAGCGGAGCTGAGAGAGCCCTACTAG 636
Db 601 AAGCCCAACGAGCGGAGCTGAGAGAGCCCTACTAG 636

RESULT 3
AF326737
LOCUS
DEFINITION Mus musculus sclerostin (Sost) gene, complete cds; and Meox1 gene,
partial sequence.
ACCESSION AF326737
VERSION AF326737.2 GI:20150447
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 81806)
Brunkow,M.E., Gardner,J.C., Van Ness,J., Paepert,B.W.,
Kovacevich,B.R., Prolli,S., Skonier,J.E., Zhao,L., Sabo,P.J.,
Fu,Y.H., Altsch,R.S., Gillett,L., Colbert,T., Tacconi,P., Galas,D.,
Hamerma,H., Beighton,P. and Mulligan,J.T.
Bone dysplasia sclerosteosis results from loss of the SOST gene
product, a novel cysteine knot-containing protein
Am. J. Hum. Genet. 68 (3), 577-589 (2001)
21090529
11179006

2 (bases 1 to 81806)
Staebling-Hampton,K., Prolli,S., Paepert,B.W., Zhao,L., Charmley,P.,
Brown,A., Gardner,J.C., Galas,D., Schatzman,R.C., Beighton,P.,
Papapoulos,S., Hamerma,H. and Brunkow,M.E.
A 52 kb deletion in the SOST - MEKX1 intergenic region on 17q12-q21
is associated with van Buchem disease in the Dutch population
Unpublished
3 (bases 1 to 17423)
Brunkow,M.E., Gardner,J.C., Van Ness,J., Paepert,B.W.,
Kovacevich,B.R., Prolli,S., Skonier,J.E., Zhao,L., Sabo,P.J.,
Fu,Y.H., Altsch,R.S., Gillett,L., Colbert,T., Tacconi,P., Galas,D.,
Hamerma,H., Beighton,P. and Mulligan,J.T.
Direct Submission
Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631
220th St. SE, Bothell, WA 98021, USA
4 (bases 1 to 81806)
Staebling-Hampton,K., Prolli,S., Paepert,B.W., Zhao,L., Charmley,P.,
Brown,A., Gardner,J.C., Galas,D., Schatzman,R.C., Beighton,P.,
Papapoulos,S., Hamerma,H. and Brunkow,M.E.
Direct Submission
Submitted (02-AUG-2001) Genomics, Celltech R&D Inc., 1631 220th St
SE, Bothell, WA 98021, USA
Sequence update by submitter
On Apr 15, 2002 this sequence version replaced gi:13161013.
Location/Qualifiers
1. 81806
/organism="Mus musculus"
/strain="129SV/J"
/db_xref="taxon:10090"
/chromosome="11"
/map="between Sost and Meox1"
<8036..>11163
/gene="Sost"
join(<8036..8249,10742..>11163)
/gene="Sost"
/product="sclerostin"
join(8036..8249,10742..11163)
/gene="Sost"
/codon_start=1
/product="sclerostin"
/protein_id="AAK13452.1"
/db_xref="GI:13161014"
/translation="MPSIAPLCILVHAAPCAVEGGQMAFNDATVEVPIGGEYP
EPPENNQTMRAENGGRPHNRYDAKVSEISCEMLHYTFLLDGPCKSAKPYTELV

TITLE
JOURNAL
REMARK
COMMENT
FEATURES
source

gene
mrna
CDS

CSGCCPALLPNAIGRVKMMRPNQDPFRCIIPDRYARVOLLCPGAGAPRSRKYRLV
ASCKCKRLTRFHNOSELKDFPETFARPOKXKPRRGARAGANAELENAI"
81492..>81806
/gene="Meox1"
81492..>81806
/gene="Meox1"
/product="Meox1"
BASE COUNT 20982 a 20323 c 20010 g 20491 t
ORIGIN

Query Match 66.7%; Score 424; DB 10; Length 81806;
Best Local Similarity 100.0%; Pred. No. 8.9e-200;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 213 AGATGTGTCGAGTACAGTGTGCGGAGCTGACATACACCCGCTTCTGACAGAGCGCC 272
Db 10740 AGATGTGTCGAGTACAGTGTGCGGAGCTGACATACACCCGCTTCTGACAGAGCGCC 10799

Qy 273 ATGCCGACGCGCAAGCGCGTACACGAGTTGTGTCTCCGACAGTGGCGCCCGCGC 332
Db 10800 ATGCCGACGCGCAAGCGCGTACACGAGTTGTGTCTCTCCGACAGTGGCGCCCGCGC 10859

Qy 333 GCTGCTGCCCAACGCCATCGCGCGGTGAGTGTGCGCGCCGAAACGACCGGATTTCG 392
Db 10860 GCTGCTGCCCAACGCCATCGCGCGGTGAGTGTGCGCGCCGAAACGACCGGATTTCG 10919

Qy 393 CTGCATCCCGGATCGCTACCGCGCGGACGCGGTGACGCTGTCTGTCCCGGCGCGCGC 452
Db 10920 CTGCATCCCGGATCGCTACCGCGCGGACGCGGTGACGCTGTCTGTCCCGGCGCGCGC 10979

Qy 453 GCGCGCTGCGCAAGTGTGCTGTGAGCTCTGTCGCAAGTGCAGACCGCTCACCCGCTT 512
Db 10980 GCGCGCTGCGCAAGTGTGCTGTGAGCTCTGTCGCAAGTGCAGACCGCTCACCCGCTT 11039

Qy 513 CCACACACAGTGTGAGCTTCAAGACTTTCGGCGCGAGACCGCGCGCGCGCAAGAGGTG 572
Db 11040 CCACACACAGTGTGAGCTTCAAGACTTTCGGCGCGAGACCGCGCGCGCGCAAGAGGTG 11099

Qy 573 CAAGCG 632
Db 11100 CAAGCG 11159

Qy 633 CTAG 636
Db 11160 CTAG 11163

RESULT 4
AC068782.2
WPCOMMENT
Sequence split into 5 fragments LOCUS AC068782 Accession AC068782
Fragment Name Begin End
AC068782_0 1 110000
AC068782_1 100001 210000
AC068782_2 200001 310000
AC068782_3 300001 410000
AC068782_4 400001 441130
Continuation (3 of 5) of AC068782 from base 200001 (AC068782 Mus musculus chromosome 11 c

Query Match 66.7%; Score 424; DB 2; Length 110000;
Best Local Similarity 100.0%; Pred. No. 8.5e-200;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 213 AGATGTGTCGAGTACAGTGTGCGGAGCTGACATACACCCGCTTCTGACAGAGCGCC 272
Db 56600 AGATGTGTCGAGTACAGTGTGCGGAGCTGACATACACCCGCTTCTGACAGAGCGCC 56659

Qy 273 ATGCCGACGCGCAAGCGCGTACACGAGTTGTGTCTCCGACAGTGGCGCCCGCGC 332
Db 56660 ATGCCGACGCGCAAGCGCGTACACGAGTTGTGTCTCCGACAGTGGCGCCCGCGC 56719

Qy 333 GCTGCTGCCCAACGCCATCGCGCGGTGAGTGTGCGCGCCGAAACGACCGGATTTCG 392
Db 56660 GCTGCTGCCCAACGCCATCGCGCGGTGAGTGTGCGCGCCGAAACGACCGGATTTCG 56719

	Db	56720	GAGGTGGCCAAACGCGCATCGAGCGGTAGTGAAGTGTCGCCGCCGAACGGACGGATTTCCG	56779
OY	393	CTGCATCCCGGATCGCTACCCGCGCAGCGGGTGCAGCTGCTGTTCCCCCGGGCGCGCGC	452	
Db	56780	CTGCATCCCGGATCCCTACCCGCGCCAGCGGGTGTGACGCTGTTCCCCCGGGCGCGCGC	56839	
OY	453	GCCGCGCTCGSCCAAGTGCCTGTCTGTGGCCCTGTGCAAGTGCAGAAGCGCTCACCCGCTT	512	
Db	56840	GCCGCGCTCGSCCAAGTGCCTGTCTGTGGCCCTGTGCAAGTGCAGAAGCGCTCACCCGCTT	56899	
OY	513	CCACAACCACTGCGAGCTCAGGACTTGGGGCCCGAGAACCCGCGCGCCGAGAAAGGCTCG	572	
Db	56900	CCACAACCACTGCGAGCTCAGGACTTGGGGCCCGAGAACCCGCGCGCGCAGAAAGGCTCG	56959	
OY	573	CNAAGCCGCGCGCCCGCGCGCCCGGAGACCNAAGCCAACAGCGGAGACTTGAGAAAGCTTA	632	
Db	56960	CNAAGCCGCGCGCCCGCGCGCCCGGAGACCNAAGCCAACAGCGGAGACTTGAGAAAGCTTA	57019	
OY	633	CTAG 636		
Db	57020	CTAG 57023		
RESULT 5				
LOCUS	AC012296	205277 bp	DNA	linear HTG 16-AUG-2002
DEFINITION	Mus musculus chromosome 11 clone RP23-346P7 map 11, *** SEQUENCING IN PROGRESS ***, 10 unordered pieces.			
ACCESSION	AC012296	GI:2267737		
VERSION	AC012296.10			
KEYWORDS	HTG; HTGS_PHASE1; HTGS_FULUTOP; HTGS_ACTIVERIN.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
AUTHORS	Birren, B., Nusbaum, C. and Lander, E.			
TITLE	Mus musculus chromosome 11, clone RP23-346P7			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 205277)			
AUTHORS	Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Balwin, J., Bartha, N., Beckwith, R., Boguski, L., Bouckheiser, B., Brown, A., Castle, A., Collins, M., Collins, S., Collymore, A., Cooke, P., DeBella, K., Dewar, K., Domino, M., Donnelly, L., Doyle, M., Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D., Horton, L., Howland, J., Gardyna, S., Grant, G., Hagos, B., Heath, A., Klein, J., Lech, J., Liu, C., Johnson, R., Jones, C., Kahn, L., Karas, A., Klein, J., Lech, J., Liu, C., Locke, K., MacDonald, P., Margulis, N., McEwan, P., McGuire, A., McKernan, K., McDonald, J., Meldrum, J., Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talmas, J., Testa, S., Tirelli, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X., Westad, D., Ye, W.J., Zimmer, A. and Zody, M.			
TITLE	Direct Submission			
JOURNAL	Submitted (22-OCT-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA			
REFERENCE	3 (bases 1 to 205277)			
AUTHORS	Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Bartha, N., Bastien, V., Bloom, F., Boguski, L., Bouckheiser, B., Camarero, J., Chang, J., Chazaro, T., Choquet, J., Collymore, A., Cook, A., Cooke, P., DeBella, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., MacDonald, P., Major, J., Matthews, C., McCarthy, M., Meldrum, J., Menusz, L., Mohova, T., Miura, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Nord, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhhang, P., Pierre, N., Raymond, C., Retra, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schnapack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talmas, J.,			

TITLE
JOURNAL
COMMENT

Testfve, S., Theodore, J., Tophan, K., Travers, M., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., ...
 Zemek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (16-AUG-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Aug 16, 2002 this sequence version replaced gi:20177756.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RW/RepeatMasker.html>
 ----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: MMRB
 Web site: <http://www.seq.wi.mit.edu>
 Contact: submissions@genome.wi.mit.edu
 ----- Project Information -----
 Center project name: U205
 Center clone name: 346_P_7

* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

*	92178	92178: contig of 92178 bp in length
*	92179	92179: gap of 100 bp
*	92219	92219: contig of 36426 bp in length
*	92279	92279: gap of 100 bp
*	128705	128705: contig of 11468 bp in length
*	128805	140272: contig of 100 bp
*	140273	140273: gap of 100 bp
*	140373	147001: contig of 6629 bp in length
*	147002	147101: gap of 100 bp
*	147102	1701176: contig of 23075 bp in length
*	170217	1702176: gap of 100 bp
*	1702177	173609: contig of 3333 bp in length
*	173610	173709: gap of 100 bp
*	173710	175712: contig of 2013 bp in length
*	175723	175832: gap of 100 bp
*	175823	181949: contig of 6127 bp in length
*	181950	182049: gap of 100 bp
*	182050	191157: contig of 9108 bp in length
*	191258	191257: gap of 100 bp
*	191258	205277: contig of 14020 bp in length

FEATURES
source

```

/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="11"
/map="11"
/clone="RP23-346P7"
/clone_1lb="RPC1-23 Female Mouse BAC"
BASE COUNT      53425 a 49464 c 50390 g 50980 t 1018 otheers
ORIGIN

```

Query Match	66.7%;	Score 424;	DB 2;	Length 205277;
Best Local Similarity	100.0%;	Pred. No. 7.8e-200;		
Matches 424;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

Qy	213	AGATGTCGCCGATACAGTGTCCCGGACCTGCATTAACCCGGTTCCTGACAAAGGGCC	272
Db	57866	AGATGTGTCCGATACAGTGTCCCGGACCTGCATTAACCCGGTTCCTGACAAAGGGCC	57925
Qy	273	ATGCCGAGCGCCAAACCGGTACCCAGATGTGTGTCGCGCCAGTGCAGGCCCCCGCGG	332
Db	57926	ATGCCGAGCGCCAAACCGGTACCCAGATGTGTGTCGCGCCAGTGCAGGCCCCCGCGG	57985
Qy	333	GTCGTCGCCCAACGCATCGGCGCGGTGAAGTGTGGCGCCCGAACGACCGGATTTTCG	392
Db	57986	GTCGTCGCCCAACGCATCGGCGCGGTGAAGTGTGGCGCCCGAACGACCGGATTTTCG	58045

Qy 393 CTGATCCCGGATGCTACCGCGCGGAGGAGTGTGCTGTGCTCCCGGCGCGCGC 452
 Db 58046 CTGATCCCGGATGCTACCGCGCGGAGGAGTGTGCTGTGCTCCCGGCGCGCGC 58105
 Qy 453 GCGCGCTCGGCGAAGGAGTGTGAGGCTGTGCAAGGCAAGGCGCTCAACCGCCTT 512
 Db 58106 GCGCGCTCGGCGAAGGAGTGTGAGGCTGTGCAAGGCAAGGCGCTCAACCGCCTT 58165
 Qy 513 CCAACACAGTCGAGGCTCAAGGACTTCGGGCGCGAGACCGCGCGCGCGAGAGGTCG 572
 Db 58166 CCAACACAGTCGAGGCTCAAGGACTTCGGGCGCGAGACCGCGCGCGCGAGAGGTCG 58225
 Qy 573 CAAGCCCGCGCGCGCGCGCGCGGAGCCAAAGCCAAACAGCGGAGTGGAGAACGCTTA 632
 Db 58226 CAAGCCCGCGCGCGCGCGCGGAGCCAAAGCCAAACAGCGGAGTGGAGAACGCTTA 58285
 Qy 633 CTAG 636
 Db 58286 CTAG 58289

RESULT 6
 AC068807 208135 bp DNA linear ROD 24-JUL-2001
 LOCUS Mus musculus 11 BAC RP23-346P7 (Rosewell Park Cancer Institute Human
 DEFINITION BAC Library) complete sequence.
 AC068807
 VERSION AC068807.21 GI:15004885
 KEYWORDS HTG.
 SOURCE Mus musculus.
 ORGANISM Mus musculus.

REFERENCE
 AUTHORS Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mms.
 1 (bases 1 to 208135)
 Metzger, M.L., Lewis, L.R., Hume, J., Edwards, C., Harris, C.,
 Dederich, D., Thomas, S., Okumura, G., Carlock, C., Garner, T.,
 Addison, S., Pace, A., Williams, G., Bonin, D., Brooks, A., Brown, J.,
 Buha, C., Bunac, C., Burkett, C., Chacko, J., Chen, G., Chen, Z.,
 Cox, C., Davis, C., Delgado, O., Ding, Y., Dugan-Rocha, S.,
 Fernandez, C., Ferraguto, D., Forcum-Tamsey, J., Gill, R.,
 Gorrell, J.H., Gunatane, P., Haller, G., Hernandez, J., Hognes, M.,
 Hosak, H., Hou, X., Huber, J., Jackson, L., Jia, Y., Kelly, J., Kelly, S.,
 Kovar, C., Liu, J., Liu, W., Louie, H., Lozano, R.J., Martin, R.,
 Massey, E., McLeod, M.P., Mei, G., Moore, S., Morgan, M., Morris, B.,
 Neal, D., Nelson, A., Nguyen, R., Nguyen, N., Ogih, M., Parish, B.,
 Perez, L., Reiter, D., Say, J., Shen, H., Vasquez, L., Watlington, S.,
 Williamson, A., Wrenford, G., Zhou, X., Bouck, J., Hodgson, A.,
 Muzny, D.M., Rives, M., Scherer, S., Sodergren, E., Weinstein, G.,
 Worley, K., and Gibbs, R.
 Direct Submission
 Unpublished
 2 (bases 1 to 208135)
 Worley, K.C.
 Direct Submission
 Submitted (10-MAY-2000) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 208135)
 Worley, K.C.
 Direct Submission
 Submitted (24-JUL-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Jul 24, 2001 this sequence version replaced gi:14787157.
 Sequencing is completed to a minimum standard of double strand
 coverage with a minimum of 2 clones and 2 reads with no ambiguities
 or 2 chemistries with a minimum of 2 clones and 3 reads with no
 ambiguities; if the sequence quality does not meet this standard,
 it will be indicated in the annotation.
 The repeat regions shown were identified using RepeatMasker by
 Adrian Smit.

Sequence similarities were identified using Powerblast by Jinghui
 Zhang.
 Exon/Intron boundaries of identified genes were chosen if there
 were canonical splice junctions that maintained sequence continuity
 across the splice junctions.
 Location/Qualifiers

FEATURES
 source
 1..208135
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /chromosome="11"
 /clone="RP23-346P7"
 987..1012
 /rpt_family="AT_rich"
 complement(1028..1167)
 /rpt_family="RSINE1"
 complement(1227..1322)
 /rpt_family="PB1D10"
 2242..2356
 /rpt_family="PB1"
 2408..2433
 /rpt_family="B4"
 2434..2546
 /rpt_family="B1_MM"
 2580..2689
 /rpt_family="TAGA)n"
 2704..2730
 /rpt_family="CA)n"
 2816..3012
 /rpt_family="B4"
 3061..3117
 /rpt_family="CA)n"
 3641..3701
 /rpt_family="TC)n"
 4381..4599
 /rpt_family="B3"
 4606..4818
 /rpt_family="B2"
 5029..5068
 /rpt_family="CA)n"
 complement(5075..5241)
 /rpt_family="B3"
 complement(6601..6740)
 /rpt_family="RSINE1"
 6926..7005
 /rpt_family="B3"
 8178..8205
 /rpt_family="TG)n"
 complement(9046..9149)
 /rpt_family="PB1"
 9190..9358
 /rpt_family="MER2"
 9359..9707
 /rpt_family="OR1D"
 9708..9801
 /rpt_family="MER2"
 9802..9831
 /rpt_family="TTTG)n"
 9832..9886
 /rpt_family="MER2"
 10219..10239
 /rpt_family="TG)n"
 complement(12931..13108)
 /rpt_family="L1MC4"
 complement(13166..13231)
 /rpt_family="MIR"
 complement(14252..14426)
 /rpt_family="MIR"
 complement(14911..15037)
 /rpt_family="L2"
 15543..15794
 /rpt_family="B4"
 15849..15890

```

repeat_region      /rpt_family="(CA)n"
                    15971..16099
repeat_region      /rpt_family="L2"
                    16336..16404
repeat_region      /rpt_family="BC1_MM"
                    16419..16550
repeat_region      /rpt_family="B1-F"
                    17412..17518
repeat_region      /rpt_family="B1-F"
                    17538..17710
repeat_region      /rpt_family="B3A"
                    18267..18397
repeat_region      /rpt_family="B3A"
                    18508..18679
repeat_region      /rpt_family="B3"
                    18939..18982
repeat_region      /rpt_family="(TPAG)n"
                    18989..19122
repeat_region      /rpt_family="B1_MM"
                    20446..20476
repeat_region      /rpt_family="(TG)n"
                    20593..20889
repeat_region      /rpt_family="B4A"
                    20964..21091
repeat_region      /rpt_family="PB1D10"
                    22076..22187
STS
repeat_region      /standard_name="163518"
                    complement(22629..22743)
repeat_region      /rpt_family="MIR"
                    22748..22899
repeat_region      /rpt_family="B4A"
                    23282..23305
repeat_region      /rpt_family="(A)n"
                    complement(23322..23355)
repeat_region      /rpt_family="B2"
                    complement(23381..23525)
repeat_region      /rpt_family="B2"
                    24185..24255
repeat_region      /rpt_family="(TG)n"
                    24921..25202
repeat_region      /rpt_family="Lx6"
                    complement(25271..25346)
repeat_region      /rpt_family="B3"
                    complement(25352..25444)
repeat_region      /rpt_family="PB1D10"
                    26173..26320
repeat_region      /rpt_family="L1MD"
                    26434..26601
repeat_region      /rpt_family="L1MD"
                    26720..26801
repeat_region      /rpt_family="RSINE1"
                    26802..26906
repeat_region      /rpt_family="B4A"
                    27067..27454
repeat_region      /rpt_family="PB1D7"
                    27555..27575
repeat_region      /rpt_family="(A)n"
                    27956..28217
repeat_region      /rpt_family="B4"
                    28275..28459
repeat_region      /rpt_family="L1M4"
                    28460..28615
repeat_region      /rpt_family="B3A"
                    complement(28701..28884)
repeat_region      /rpt_family="B2"
                    28937..29077
repeat_region      /rpt_family="L1M4"
                    29241..29376
repeat_region      /rpt_family="B4A"
                    29323..29410
repeat_region      /rpt_family="PB1D10"

```

```

repeat_region      29438..29691
repeat_region      /rpt_family="B4A"
                    29696..29823
repeat_region      /rpt_family="(CA)n"

Query Match
Best Local Similarity 100.0%; Pred. No. 7.8e-200;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 213 AGATGTGCCAGTACAGCTCCCGGAGCTGCACTACACCCGCTTCTGTACACAGCGCC 272
Db 57847 AGATGTGCCAGTACAGCTCCCGGAGCTGCACTACACCCGCTTCTGTACACAGCGCC 57906
QY 273 ATGCCGACGCGCCAAAGCCGCGGACGAGTTGTGTCTCCGCGGACGAGCCCGCGCG 332
Db 57907 ATGCCGACGCGCCAAAGCCGCGGACGAGTTGTGTCTCCGCGGACGAGCCCGCGCG 57966
QY 333 GCTGCTGCCCAACGCGCATCGGCGCGGTGAAGTGTGCGCGCCGAAACGACCGGATTCCG 392
Db 57967 GCTGCTGCCCAACGCGCATCGGCGCGGTGAAGTGTGCGCGCCGAAACGACCGGATTCCG 58026
QY 393 CTGCATCCCGGATGCTACCCGCGGACGCGGTGCTGTGCTGCTCCCGGCGCGCGC 452
Db 58027 CTGCATCCCGGATGCTACCCGCGGACGCGGTGCTGTGCTGCTCCCGGCGCGCGC 58086
QY 453 GCGCGCTCGCGCAAGGCGCTGTGTGCTGTGCAAGTGCAGAGGCGCTCACCCGCTT 512
Db 58087 GCGCGCTCGCGCAAGGCGCTGTGTGCTGTGCAAGTGCAGAGGCGCTCACCCGCTT 58146
QY 513 CCACAACGATCGAGCTCAAGGACTTTCGCGCGGAGACGCGCGCGCGCAAGAGGCTCG 572
Db 58147 CCACAACGATCGAGCTCAAGGACTTTCGCGCGGAGACGCGCGCGCGCAAGAGGCTCG 58206
QY 573 CAAGCGCGCGCGCGCGCGCGCGGAGCCAAAGCCAAACGAGCGGAGTGAAGAGCGCTA 632
Db 58207 CAAGCGCGCGCGCGCGCGCGGAGCCAAAGCCAAACGAGCGGAGTGAAGAGCGCTA 58266
QY 633 CTAG 636
Db 58267 CTAG 58270

RESULT 7
ALS91145
LOCUS
DEFINITION Mouse DNA sequence from clone Rp23-3987 on chromosome 11, complete
sequence.
ACCESSION ALS91145
VERSION ALS91145.24 GI:22474402
KEYWORDS HTG.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE
AUTHORS Garner,P.
TITLE Direct Submmission
JOURNAL Submitted (22-AUG-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Aug 23, 2002 this sequence version replaced gi:2203099.
COMMENT
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
-----
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.

```

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information on the WormPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-398F7 is from the RPI-23 Mouse PAC Library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBAC3.6.

FEATURES
source
Location/Qualifiers
1..198508
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-398F7"
/clone_1b="RPI-23"

BASE COUNT 48902 a 47818 c 50456 g 51332 t
ORIGIN

Query Match 58.6%; Score 373; DB 10; Length 198508;
Best Local Similarity 99.8%; Pred. No. 1.7e-174;
Matches 423; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 213 AGATTGTGTCGAGTACAGCTGCGCGGAGCTACACCCGCTTCTTACAGAGCGGCC 272
Db 187744 AGATTGTGTCGAGTACAGCTGCGCGGAGCTACACCCGCTTCTTACAGAGCGGCC 187803
QY 273 ATCCCGCAGCGCCAGCGGCTGACCGAGTGTGTGCTCCGAGTGGGGCGCCGCGCG 332
Db 187804 ATCCCGCAGCGCCAGCGGCTGACCGAGTGTGTGCTCCGAGTGGGGCGCCGCGCG 187863
QY 333 GCTGCTGCGCCAGCGGCTGAGTGTGTGCTCCGAGTGGGGCGCCGCGAGATTTCG 392
Db 187864 GCTGCTGCGCCAGCGGCTGAGTGTGTGCTCCGAGTGGGGCGCCGAGATTTCG 187923
QY 393 CTGCATCCCGGAGTCTTACCGCGCGGAGCTGCTGTGCTCCGCGGGCGCGCG 452
Db 187924 CTGCATCCCGGAGTCTTACCGCGCGGAGCTGCTGTGCTCCGCGGGCGCGCG 187983
QY 453 GCGCGCTCGCGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 512
Db 187984 GCGCGCTCGCGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 188043
QY 513 CCACACCACTGAGCTTCAAGACTTCCGCGCGGAGACCGCGCGCGCGAGAGGCTG 572
Db 188044 CCACACCACTGAGCTTCAAGACTTCCGCGCGGAGACCGCGCGCGCGAGAGGCTG 188103
QY 573 CAGCG 632
Db 188104 CAGCG 188163
QY 633 CTAG 636
Db 188164 CTAG 188167

RESULT 8
AC023810/c 51575 bp DNA linear HTG 04-NOV-2000
LOCUS
DEFINITION Mus musculus chromosome 11 clone RP23-252B10, *** SEQUENCING IN
PROGRESS ***; 35 unordered pieces.
AC023810
AC023810
AC023810.6 GI:11079371
KEYWORDS HTG; HTGS PHASE1.
SOURCE Mus musculus.
ORGANISM Mus musculus

REFERENCE AUTHORS

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 51575)
Metzker, M.L., Lewis, L.R., Hume, J., Edwards, C., Harris, C., Dederich, D., Thomas, S., Okwuonu, G., Carlick, C., Garner, T., Addison, S., Pace, A., Williams, G., Bonnin, D., Brooks, A., Brown, J., Bunay, C., Bunac, C., Burkett, C., Chacko, J., Chen, G., Chen, Z., Cox, C., Davis, C., Delgado, O., Ding, Y., Dugan-Rocha, S., Fernandez, C., Ferraruto, D., Forcum-Taney, J., Gill, R., Gorrell, J.H., Gunaratne, P., Haller, G., Hernandez, J., Hogues, M., Hosak, H., Hou, X., Huber, J., Jackson, L., Jia, Y., Kelly, J., Kelly, S., Kovar, C., Liu, J., Liu, W., Louisse, H., Lozano, R., J., Martin, R., Massey, E., McLeod, M.P., Mei, G., Moore, S., Morgan, M., Morris, S., Neal, D., Nelson, A., Nguyen, R., Nguyen, N., Ogund, M., Parish, B., Perez, L., Reiter, D., Say, J., Shen, H., Vasquez, L., Watlington, S., Williamson, A., Wrensford, G., Zhou, X., Bouck, J., Hodgson, A., Murry, D.M., Rives, M., Scherer, S., Sodergren, E., Weinert, G., Worley, K., and Gibbs, R.
Unpublished
Direct Submission
2 (bases 1 to 51575)
Worley, K.C.
Direct Submission
Submitted (18-FEB-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 3, 2000 this sequence version replaced gi:8248603.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: MABO
Center clone name: RP23-252B10
----- Summary Statistics
Sequencing vector: M13; L08821
Chemistry: Dye-terminator Big Dye 7% of reads
Chemistry: Dye-terminator Big Dye 24% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 42708 bases at least Q40
Consensus quality: 59319 bases at least Q30
Consensus quality: 66755 bases at least Q20
Estimated insert size: 39143; sum-of-coverage estimation
Quality coverage: 0x in Q20 bases; agarose-fp estimation
Quality coverage: 0.4x in Q20 bases; sum-of-coverage estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbankdraft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 35 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
2683: contig of 2682 bp in length
2783: gap of unknown length
2783: contig of 3111 bp in length
5894: gap of unknown length
5993: gap of 1446 bp in length
7439: contig of 1446 bp in length
7539: gap of unknown length
7540: contig of 1767 bp in length
9306: gap of unknown length
9307: contig of 1856 bp in length
11263: gap of unknown length
11263: contig of 1374 bp in length
11363: gap of unknown length
12737: gap of unknown length
12836: contig of 1116 bp in length
13952: gap of unknown length
13953: contig of 1154 bp in length
14053: gap of unknown length
15207: gap of unknown length
15306: gap of unknown length

```

* 15307 16973: contig of 1667 bp in length
* 16974 17073: gap of unknown length
* 17074 18081: contig of 1008 bp in length
* 18082 18181: gap of unknown length
* 18182 19630: contig of 1449 bp in length
* 19631 19730: gap of unknown length
* 19731 20732: contig of 1002 bp in length
* 20733 20832: gap of unknown length
* 20833 22555: contig of 1723 bp in length
* 22556 22656: gap of unknown length
* 22657 23846: contig of 1191 bp in length
* 23847 23947: gap of unknown length
* 23948 25489: contig of 1543 bp in length
* 25490 25590: gap of unknown length
* 25591 26913: contig of 1324 bp in length
* 26914 27013: gap of unknown length
* 27014 28158: contig of 1145 bp in length
* 28159 28258: gap of unknown length
* 28259 29424: contig of 1166 bp in length
* 29425 29525: gap of unknown length
* 29526 30594: contig of 1070 bp in length
* 30595 30695: gap of unknown length
* 30696 32013: contig of 1319 bp in length
* 32014 32114: gap of unknown length
* 32115 33542: contig of 1429 bp in length
* 33543 33642: gap of unknown length
* 33643 34705: contig of 1063 bp in length
* 34706 34805: gap of unknown length
* 34806 35911: contig of 1106 bp in length
* 35912 36011: gap of unknown length
* 36012 37239: contig of 1228 bp in length
* 37240 37339: gap of unknown length
* 37340 38613: contig of 1274 bp in length
* 38614 38713: gap of unknown length
* 38714 39885: contig of 1172 bp in length
* 39886 39985: gap of unknown length
* 39986 41336: contig of 1351 bp in length
* 41337 41436: gap of unknown length
* 41437 42561: contig of 1125 bp in length
* 42562 42661: gap of unknown length
* 42662 43685: contig of 1024 bp in length
* 43686 43785: gap of unknown length
* 43786 45017: contig of 1232 bp in length
* 45018 45117: gap of unknown length
* 45118 46507: contig of 1399 bp in length
* 46508 46607: gap of unknown length
* 46608 47781: contig of 1174 bp in length
* 47782 47881: gap of unknown length
* 47882 49024: contig of 1143 bp in length
* 49025 49124: gap of unknown length
* 49125 50392: contig of 1268 bp in length
* 50393 50492: gap of unknown length
* 50493 51575: contig of 1083 bp in length.

```

FEATURES
source
1. 51575
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-252B10"

BASE COUNT 12565 a 11207 c 11347 g 13002 t 3454 others
ORIGIN

Query Match 33.6% Score 214; DB 2; Length 51575;
Best Local Similarity 100.0%; Pred. No. 2.1e-95;
Matches 214; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGAGCCCTACTAGCCCGTGCCTCATCTGCTACTTGTGACGCTGCTTGTGCT 60
DB 8375 ATGAGCCCTACTAGCCCGTGCCTCATCTGCTACTTGTGACGCTGCTTGTGCT 8316
QY 61 GTGAGGCGCCAGGGGTGGCAAGCTTCAGAAATGATGCCACAGAGTTCATCCAGGGCTT 120
DB 8315 GTGAGGCGCCAGGGGTGGCAAGCTTCAGAAATGATGCCACAGAGTTCATCCAGGGCTT 8256

```

```

QY 121 GGAGAGTACCCCGAGCCTCTCTCTGAGAGCAACAGACCATGAACCGGGCGAGGAATGGA 180
DB 8255 GGAGAGTACCCCGAGCCTCTCTCTGAGAGCAACAGACCATGAACCGGGCGAGGAATGGA 8196
QY 181 GGCAGACCTCCCGACCATCCCTATGACGCCAAG 214
DB 8195 GGCAGACCTCCCGACCATCCCTATGACGCCAAG 8162

```

RESULT 9
AF326741 674 bp mRNA linear ROD 28-FEB-2001
LOCUS
DEFINITION Rattus norvegicus sclerostin mRNA, complete cds.
ACCESSION AF326741
VERSION AF326741.1 GI:13161025
KEYWORDS
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
AUTHORS 1 (bases 1 to 674)
Brunkow, M.E., Gardner, J.C., Van Ness, J., Paepker, B.W.,
Kovacevich, B.R., Prohl, S., Skonier, J.E., Zhao, L., Sabo, P.J.,
Fu, Y.H., Altsch, R.S., Gillett, L., Colbert, T., Tacconi, P., Galas, D.,
Hamersma, H., Beighon, P. and Mulligan, J.T.
Bone dysplasia sclerosteosis results from loss of the SOST gene
product, a novel cystine knot-containing protein
Am. J. Hum. Genet. 68 (3), 577-589 (2001)
MEDLINE 21090529
PUBMED 11179006

TITLE
JOURNAL 2 (bases 1 to 674)
PUBMED 11179006
REFERENCE
AUTHORS Brunkow, M.E., Gardner, J.C., Van Ness, J., Paepker, B.W.,
Kovacevich, B.R., Prohl, S., Skonier, J.E., Zhao, L., Sabo, P.J.,
Fu, Y.H., Altsch, R.S., Gillett, L., Colbert, T., Tacconi, P., Galas, D.,
Hamersma, H., Beighon, P. and Mulligan, J.T.
Direct Submission
Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631
220th St. SE, Bothell, WA 98021, USA

FEATURES
source
1. 674
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"

5'UTR
CDS
1..32
33..674
/codon_start=1
/product="sclerostin"
/protein_id="AAK13456.1"
/db_xref="GI:13161026"

BASE COUNT 128 a 240 c 208 g 98 t
ORIGIN

Query Match 16.4% Score 104; DB 10; Length 674;
Best Local Similarity 99.4%; Pred. No. 1.8e-40;
Matches 154; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 388 TTCGGCTGATCCCGATCGGTACCGCGCGAGGCGGAGTGTGTGCGCCGCGGCGC 447
DB 426 TTCGGCTGATCCCGATCGGTACCGCGCGAGGCGGAGTGTGTGCGCCGCGGCGC 485
QY 448 GCGGCGCGCGCTCGCGGAGGTGCTGTGCTGTGCAAGTCAAGCGCTTACC 507
DB 486 GCGGCGCGCGCTCGCGGAGGTGCTGTGCTGTGCAAGTCAAGCGCTTACC 545
QY 508 CGCTTCACACCACTGCGAGTCAAGACTTGG 542
DB 546 CGCTTCACACCACTGCGAGTCAAGACTTGG 580

```

Db 99 PVTGVSGGCGPARLLPNAIGRGKWRPSPGDPFRCTIDRYRAORVOLLCPGGEAPAPARK 158
 QY 136 VRLVASCCKRLLTFPHNOSLKDFTGEARPOKGRKPRPARASAKANCAELENNAY 190
 Db 159 VRLVASCCKRLLTFPHNOSLKDFTGEARPOKGRKPRPARASAKANCAELENNAY 213

RESULT 10
 AAM73067
 ID AAM73067 standard; Protein; 139 AA.

XX AAM73067;

XX 06-NOV-2001 (first entry)

XX Human bone marrow expressed probe encoded protein SEQ ID NO: 33373.

XX Human; bone marrow expressed exon; gene expression analysis; probe;
 KM microarray; cancer; leukaemia; lymphoma; myeloma.

XX Homo sapiens.

XX WO200157276-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00668.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human bone marrow -

XX Example 4; SEQ ID NO: 33373; 658bp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

XX probes which are derived from genomic sequences expressed in the human

XX bone marrow. They can be used to measure gene expression in bone marrow

XX samples, which may enable the improved diagnosis and treatment of cancers

XX such as lymphoma, leukaemia and myeloma. The present sequence is a

XX protein encoded by one of the probes of the invention.

XX Sequence 139 AA;

XX Query Match 73.2%; Score 139; DB 22; Length 139;

XX Best Local Similarity 100.0%; Pred. No. 3.4e-137;

XX Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 VSEYSGRELFHTYVVDGPRSAKPTTELVCSCGCCPARLLPAITRGKWMRSPGDPFRIC 111

Db 1 VSEYSGRELFHTYVVDGPRSAKPTTELVCSCGCCPARLLPAITRGKWMRSPGDPFRIC 60

QY 112 IPDRYRAORVOLLCPGGEAPAPARKVLLVASCCKRLLTFPHNOSLKDFTGEARPOKGRK 171

Db 61 IPDRYRAORVOLLCPGGEAPAPARKVLLVASCCKRLLTFPHNOSLKDFTGEARPOKGRK 120

QY 172 PRRARSASAKANCAELENNAY 190

Db 121 PRRARSASAKANCAELENNAY 139

RESULT 11
 ABG42911
 ID ABG42911 standard; Peptide; 139 AA.

XX ABG42911;

XX 19-AUG-2002 (first entry)

XX Human peptide encoded by genome-derived single exon probe SEQ ID 32576.

XX Human; single exon probe; asthma; lung cancer; COPD; ILD;

XX chronic obstructive pulmonary disease; interstitial lung disease;

XX familial idiopathic pulmonary fibrosis; neurofibromatosis;

XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;

XX Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemorrhage;

XX pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;

XX pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;

XX primary ciliary dyskinesia; pulmonary hypertension;

XX hyaline membrane disease.

XX Homo sapiens.

XX WO200186003-A2.

XX 15-NOV-2001.

XX 30-JAN-2001; 2001WO-US00665.

XX 04-FEB-2000; 2000US-180312P.

XX 26-MAY-2000; 2000US-207456P.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-234687P.

XX 27-SEP-2000; 2000US-236359P.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2002-114183/15.

XX Spatially-addressable set of single exon nucleic acid probes, used to

XX measure gene expression in human lung samples -

XX Claim 27; SEQ ID NO 32576; 634bp; English.

XX The invention relates to a spatially-addressable set of single exon

XX nucleic acid probes for measuring gene expression in a sample derived

XX from human lung comprising single exon nucleic acid probes having one of

XX 12614 nucleic acid sequences mentioned in the specification, or their

XX complements or the 12387 open reading frames derived from the 12614

XX probes. Also included are a microarray comprising the novel set of

XX probes; the novel set of probes which hybridize at high stringency to a

XX nucleic acid expressed in the human lung; measuring gene expression in a

XX sample derived from human lung, comprising (a) contacting the array with

XX a collection of detectably labeled nucleic acids derived from human lung

XX mRNA, and (b) measuring the label detectably bound to each probe of

XX the array; identifying exons in a eukaryotic genome, comprising

XX (a) algorithmically predicting at least one exon from genomic sequences

XX of the eukaryote; and (b) detecting specific hybridisation of detectably

XX labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,

XX having a fragment identical to the predicted exon, the probe is included

XX in the above mentioned microarray; assigning exons to a single gene,

XX comprising (a) identifying exons from genomic sequence by the method

XX above and (b) measuring the expression of each of the exons in several

XX tissues and/or cell types using hybridisation to a single exon

XX microarrays having a probe with the exon, where a common pattern of

XX expression of the exons in the tissues and/or cell types indicates that

XX the exons should be assigned to a single gene; a peptide comprising one

XX of 12011 sequences, mentioned in the specification, or encoded by the

XX probes/open reading frames (ORF). The probes are used for gene

XX expression analysis, and for identifying exons in a gene, particularly

CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karagenen syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a peptide/protein
 CC encoded by a single exon probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 139 AA;

Query Match 73.2%; Score 139; DB 23; Length 139;
 Best Local Similarity 100.0%; Pred. No. 3.4e-137;
 Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 VSEYSCRELHFTRYVTDGPCRSAPVTELVCSGCGPARLLPNAIGRGKWRSPGDFRC 111
 DB 1 VSEYSCRELHFTRYVTDGPCRSAPVTELVCSGCGPARLLPNAIGRGKWRSPGDFRC 60
 QY 112 IPDRYRAQRYVOLLCPGGEAPRARKVRLVASCKCKRLTRFHNOSLKDFTGARPOKGRK 171
 DB 61 IPDRYRAQRYVOLLCPGGEAPRARKVRLVASCKCKRLTRFHNOSLKDFTGARPOKGRK 120
 QY 172 PRPARSAKANQAELFNAY 190
 DB 121 PRPARSAKANQAELFNAY 139

RESULT 12
 ID AAB26105 standard; Protein; 367 AA.

XX AAB26105;

DT 15-JAN-2001 (first entry)

XX Human DAN/Cerberus-related protein 6 (hDCR6) #1.

KM Human; DNA/Cerberus-related protein 6; hDCR6; morphogenic protein;
 KM antagonist; BMP; cell growth; cell differentiation; bone formation;
 KM gene therapy.

XX Homo sapiens.

XX W0200055193-A2.

XX 21-SEP-2000.

XX 02-MAR-2000; 2000WO-US05537.

XX 12-MAR-1999; 99US-0124118.

XX (REGG-) REGENERON PHARM INC.

XX Economides AM;

XX WPI; 2000-638179/61.

XX N-PSDB; AAA94049, AAA94050.

XX Novel isolated, human DNA/Cerberus related protein 6 which include
 PT natural homologue, and polypeptides comprising DCR6 domain and nucleic
 PT acids encoding the proteins which are useful as probes and primers -
 XX Claim 7; Fig 2; 40pp; English.

XX The present sequence comprises the human DAN/Cerberus-related protein 6
 CC (hDCR6). Its coding sequence was isolated from a genomic DNA clone

CC following identification using computer-based 'virtual cloning'. hDCR6
 CC is closely related to the DAN and DCR5 proteins, both of which act as
 CC antagonists of morphogenic proteins such as BMP. It is possible that
 CC the hDCR6 gene and protein can be used as immunogens, modulators of cell
 CC function, growth and differentiation, to reduce undesirable bone
 CC formation, to identify DCR6 binding agents, in diagnosis, and in gene
 CC therapy.

XX Sequence 367 AA;

Query Match 73.2%; Score 139; DB 21; Length 367;
 Best Local Similarity 100.0%; Pred. No. 8e-137;
 Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 VSEYSCRELHFTRYVTDGPCRSAPVTELVCSGCGPARLLPNAIGRGKWRSPGDFRC 111
 DB 229 VSEYSCRELHFTRYVTDGPCRSAPVTELVCSGCGPARLLPNAIGRGKWRSPGDFRC 288
 QY 112 IPDRYRAQRYVOLLCPGGEAPRARKVRLVASCKCKRLTRFHNOSLKDFTGARPOKGRK 171
 DB 289 IPDRYRAQRYVOLLCPGGEAPRARKVRLVASCKCKRLTRFHNOSLKDFTGARPOKGRK 348
 QY 172 PRPARSAKANQAELFNAY 190
 DB 349 PRPARSAKANQAELFNAY 367

RESULT 13
 ID AAY96431 standard; Protein; 213 AA.

XX AAY96431;

DT 12-SEP-2000 (first entry)

XX Vervet TGF-beta binding protein (BEER).

KM osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KM BEER; gene therapy; antisense therapy; fracture; bone mineralization.

XX Cercopithecus pygerythrus.

XX W0200032773-A1.

XX 08-JUN-2000.

XX 24-NOV-1999; 99WO-US27990.

XX 27-NOV-1998; 98US-0110283.

XX (DARW-) DARWIN DISCOVERY LTD.

XX Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepker BW;

XX Van Ness J, Winkler DG,

XX WPI; 2000-412321/35.

XX N-PSDB; AAA23057.

XX Nucleic acids (1) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures

XX Claim 4; Page 122-123; 162pp; English.

XX This shows a vervet transforming growth factor-beta (TGF-beta)
 CC binding protein designated VBER. The cDNA and protein may be used for
 CC prevention, treatment and diagnosis of diseases associated with
 CC inappropriate BEER expression. For example, they may be used to treat
 CC disorders associated with decreased TGF-beta BP expression. The cDNA or
 CC vectors may be administered to treat diseases by rectifying mutations or
 CC deletions in a patient's genome that affect the activity of BEER by
 CC expressing inactive proteins or to supplement the patients own production
 CC of BEER polypeptides. The nucleic acids may be used for recombinant

CC production of BEER, gene therapy, antisense therapy, as probes for
 CC diagnostic assays and for functional studies. BEER may be used to raise
 CC antibodies and for identification of BEER modulators. BEER antagonists
 CC may be used to increase bone mineral content for the treatment of
 CC disorders such as osteopenia, osteoporosis, fractures and other disorders
 CC associated with low mineral content.

XX Sequence 213 AA;

Query Match 67.4%; Score 128; DB 21; Length 213;
 Best Local Similarity 100.0%; Pred. No. 1.6e-125;
 Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OQQAQKQKATATIIIPPLGIIPEPPPELNNKTMNAENGRRPHHPFETKDVSEYSCREL 60
 DB 24 OQQAQKQKATATIIIPPLGIIPEPPPELNNKTMNAENGRRPHHPFETKDVSEYSCREL 83
 QY 61 HFTRYVTDGPCRSAPKPYTELVCSCGCGPARLLPNAIGRGKWRPSPGDFRCI 120
 DB 84 HFTRYVTDGPCRSAPKPYTELVCSCGCGPARLLPNAIGRGKWRPSPGDFRCI 143
 QY 121 VQLCPGG 128
 DB 144 VQLCPGG 151

RESULT 14

AA96434
 ID AAY96434 standard; Protein; 176 AA.

AC AAY96434;

DT 12-SEP-2000 (first entry)

DE Bovine TGF-beta binding protein (BEER).

KM osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KM BEER; gene therapy; antisense therapy; fracture; bone mineralization.

OS Bos taurus.

XX

PN W0200032773-A1.

PD 08-JUN-2000.

PF 24-NOV-1999; 99MO-US27990.

PR 27-NOV-1998; 98US-0110283.

PA (DARW-) DARWIN DISCOVERY LTD.

PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;
 PI Van Ness J, Winkler DG;

DR WPI; 2000-412321/35.

DR N-PSDB; AAA29060.

XX

PT Nucleic acids (I) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures

XX

PS Claim 7; Page 127; 162pp; English.

CC This shows a bovine transforming growth factor-beta (TGF-beta)
 CC binding protein designated bBEER. The cDNA and protein may be used for
 CC prevention, treatment and diagnosis of diseases associated with
 CC inappropriate BEER expression. For example, they may be used to treat
 CC disorders associated with decreased TGF-beta BP expression. The cDNA or
 CC vectors may be administered to treat diseases by rectifying mutations or
 CC deletions in a patient's genome that affect the activity of BEER by
 CC expressing inactive proteins or to supplement the activity of BEER by
 CC production of BEER. The nucleic acids may be used for recombinant
 CC production of BEER, gene therapy, antisense therapy, as probes for

CC diagnostic assays and for functional studies. BEER may be used to raise
 CC antibodies and for identification of BEER modulators. BEER antagonists
 CC may be used to increase bone mineral content for the treatment of
 CC disorders such as osteopenia, osteoporosis, fractures and other disorders
 CC associated with low mineral content.

XX Sequence 176 AA;

Query Match 40.0%; Score 76; DB 21; Length 176;
 Best Local Similarity 100.0%; Pred. No. 3.4e-71;
 Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 SEYSCREIHFTRYVTDGPCRSAPKPYTELVCSCGCGPARLLPNAIGRGKWRPSPGDFRCI 112
 DB 45 SEYSCREIHFTRYVTDGPCRSAPKPYTELVCSCGCGPARLLPNAIGRGKWRPSPGDFRCI 104
 QY 113 PDRYRAQRVQLCPGG 128
 DB 105 PDRYRAQRVQLCPGG 120

RESULT 15

AA96433
 ID AAY96433 standard; Protein; 213 AA.

AC AAY96433;

DT 12-SEP-2000 (first entry)

DE Rat TGF-beta binding protein (BEER).

KM osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KM BEER; gene therapy; antisense therapy; fracture; bone mineralization.

OS Rattus norvegicus.

XX

PN W0200032773-A1.

PD 08-JUN-2000.

PF 24-NOV-1999; 99MO-US27990.

PR 27-NOV-1998; 98US-0110283.

PA (DARW-) DARWIN DISCOVERY LTD.

PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;
 PI Van Ness J, Winkler DG;

DR WPI; 2000-412321/35.

DR N-PSDB; AAA29059.

XX

PT Nucleic acids (I) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures

XX

PS Claim 6; Page 125-126; 162pp; English.

CC This shows a rat transforming growth factor-beta (TGF-beta) binding
 CC protein designated rBEER. The cDNA and protein may be used for
 CC prevention, treatment and diagnosis of diseases associated with
 CC inappropriate BEER expression. For example, they may be used to treat
 CC disorders associated with decreased TGF-beta BP expression. The cDNA or
 CC vectors may be administered to treat diseases by rectifying mutations or
 CC deletions in a patient's genome that affect the activity of BEER by
 CC expressing inactive proteins or to supplement the activity of BEER by
 CC production of BEER. The nucleic acids may be used for recombinant
 CC production of BEER, gene therapy, antisense therapy, as probes for
 CC diagnostic assays and for functional studies. BEER may be used to raise
 CC antibodies and for identification of BEER modulators. BEER antagonists
 CC may be used to increase bone mineral content for the treatment of
 CC disorders such as osteopenia, osteoporosis, fractures and other disorders
 CC associated with low mineral content.

XX Sequence 213 AA;

Query Match 17.4%; Score 33; DB 21; Length 213;
 Best Local Similarity 100.0%; Pred. No. 4e-26;
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 66 VTGDCRSARPYTELVCSGCGCPARLLPNAIGR 98
 ||||||||||||||||||||||||||||||||
 Db 89 VTGDCRSARPYTELVCSGCGCPARLLPNAIGR 121

Search completed: March 28, 2003, 14:21:57
 Job time : 34.44 secs

GenCore version 5.1.4.P5_4576
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:20:22 ; Search time 14.6933 seconds
(without alignments)
1243.118 Million cell updates/sec

Title: US-09-867-274-2

Perfect score: 190

Sequence: 1 OGMDAFKNDATETIPELGEY.....KPRPRASAKANOALENNAY 190

Scoring table:

Gapop 60.0, Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size: 0

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

1: PIR 73: *
2: PIR1: *
3: PIR2: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	4.2	428	1	TVHUEK
2	8	4.2	429	2	JCA965
3	7	3.7	130	2	D83505
4	7	3.7	148	2	C83091
5	7	3.7	153	2	AE2284
6	7	3.7	183	2	B37410
7	7	3.7	197	2	G81057
8	7	3.7	197	2	A81819
9	7	3.7	296	2	A12830
10	7	3.7	305	2	S77397
11	7	3.7	311	2	E98351
12	7	3.7	312	2	D75344
13	7	3.7	321	2	B82892
14	7	3.7	328	2	S72647
15	7	3.7	341	2	AG8834
16	7	3.7	346	2	AP3323
17	7	3.7	398	1	S24802
18	7	3.7	425	1	C89753
19	7	3.7	555	2	C97612
20	7	3.7	562	2	T05758
21	7	3.7	567	2	A87394
22	7	3.7	602	2	S47880
23	7	3.7	632	1	VGNNSY
24	7	3.7	632	2	AC1547
25	7	3.7	667	2	UC7771
26	7	3.7	681	2	E82812
27	7	3.7	781	1	TVFDFR
28	7	3.7	917	1	S15885
29	7	3.7	946	2	T16297

30	7	3.7	1008	2	B95979	probable enzyme wi
31	7	3.7	1488	2	AG2136	polyketide synthas
32	7	3.7	1940	2	A59287	myosin heavy chain
33	7	3.7	6420	2	T10283	polyketide synthas
34	6	3.2	49	2	S28215	neurotoxin Tx2 - s
35	6	3.2	65	2	B25025	malx protein - Kle
36	6	3.2	83	2	S41672	tightly associated
37	6	3.2	89	2	D75271	hypothetical prote
38	6	3.2	97	1	GECH	osteocalcin precu
39	6	3.2	102	2	G85584	unknown protein en
40	6	3.2	102	2	B90734	hypothetical prote
41	6	3.2	102	2	T51524	hypothetical prote
42	6	3.2	106	2	A82560	50S ribosomal prot
43	6	3.2	108	2	T30659	hypothetical prote
44	6	3.2	108	2	D90051	hypothetical prote
45	6	3.2	110	1	R5EC22	ribosomal protein

ALIGNMENTS

RESULT 1

TVHUEK

transforming protein elk-1 - human

C/Species: Homo sapiens (man)

C/Date: 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change 18-Jun-1999

C/Accession: A41354, S54721

R/Rao, V.N.; Huebner, K.; Isobe, M.; ar-Rushdi, A.; Croce, C.M.; Reddy, E.S.P.

Science 244, 66-70, 1989

A/Title: elk, tissue-specific ets-related genes on chromosomes X and 14 near translocated

A/Reference number: A41354; MUID:89203250; PMID:2539641

A/Accession: A41354

A/Molecule type: mRNA

A/Residues: 1-428 <RNO>

A/Cross-references: GB:M25269; NID:9538208; PIDN:AA52384.1; PID:G538209

R/Gille, H.; Kortelmann, M.; Thomae, O.; Moormaw, C.; Slaughter, C.; Cobb, M.H.; Shaw, P.I

EMBO J. 14, 951-962, 1995

A/Title: ERK phosphorylation potentiates Elk-1-mediated ternary complex formation and tr

A/Reference number: S54721; MUID:95196758; PMID:7889942

A/Accession: S54721

A/Status: preliminary

A/Molecule type: protein

A/Residues: 318-328, 'XX', 331, 336-364, 380-388, 'X', 390-392, 'X', 394-400, 'XX', 403-405, 'X', 40

C/Genetics:

A/Gene: GDB:ELK1

A/Cross-references: GDB:119867; OMIM:311040

A/Map position: Xp11.2-Xp11.2

C/Superfamily: elk-1 transforming protein; ets DNA-binding domain homology

C/Keywords: DNA binding; oncogene; transforming protein

F/7-86/Domain: ets DNA-binding domain homology <ETS>

Query Match

Best Local Similarity 4.2%; Score 8; DB 1; Length 428;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 POKGRKPR 173

DB 310 POKGRKPR 317

RESULT 2

JCA965

elk1 protein - mouse

C/Species: Mus musculus (house mouse)

C/Date: 31-Dec-1996 #sequence_revision 31-Dec-1996 #text_change 16-Jul-1999

C/Accession: JCA965; I48339; I48340; S54908

R/Greyn, D.; Ung, S.; Denhez, F.; Dehem, M.; Quatennens, B.; Begue, A.; Stehelin, D.; Me

Gene 174, 185-188, 1996

A/Title: Structure and organization of the mouse elk1 gene.

A/Reference number: JCA965; MUID:97017146; PMID:8863747

A/Accession: JCA965

A/Molecule type: mRNA

A/Residues: 1-429 <GRE>

A:Cross-references: EMBL:X87257; NID:g936634; PIDN:CAA60715.1; PID:g936635
 A:Experimental source: emdYto
 R:Giovane, A.; Pintzas, A.; Maira, S.M.; Sobieszczak, P.; Wasyluk, B.
 Genes Dev. 8, 1502-1513, 1994
 A:Title: Net, a new ets transcription factor that is activated by Ras.
 A:Reference number: A53837; MUID:95047310; PMID:7958835
 A:Accession: 148339
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-429 <RES>
 A:Cross-references: EMBL:X87257; NID:g936634; PIDN:CAA60715.1; PID:g936635
 A:Accession: 148340
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 5-132, 'T', 134-224 <RE2>
 A:Cross-references: EMBL:X23693; NID:g535922; PIDN:CAA85391.1; PID:g535923
 A:Comment: This protein belongs to the subfamily of ternary complex factor (TCF) which h
 C:Genetics:
 A:Gene: elk1
 A:Introns: 70/3; 219/3; 363/3; 397/3
 C:Superfamily: elk-1 transforming protein; ets DNA-binding domain homology
 F:7-86/Domain: ets DNA-binding domain homology <ETS>
 Query Match 4.2%; Score 8; DB 2; Length 429;
 Best Local Similarity 100.0%; Pred. No. 3.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 166 POKGRKR 173
 DB 311 POKGRKR 318

RESULT 3
 D83305
 Hypothetical protein PA2722 [imported] - Pseudomonas aeruginosa (strain PA01)
 C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: D83305
 A:Reference number: A82950; MUID:20437337; PMID:10984043
 A:Accession: D83305
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-130 <STO>
 A:Cross-references: GB:AE004700; GB:AE004091; NID:g9948792; PIDN:AG06110.1; GSPDB:GN001
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA2722

Query Match 3.7%; Score 7; DB 2; Length 130;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 135 KVRIVAS 141
 DB 12 KVRIVAS 18

RESULT 4
 C83091
 Hypothetical protein PA4441 [imported] - Pseudomonas aeruginosa (strain PA01)
 C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: C83091
 A:Reference number: A53837; MUID:95047310; PMID:7958835
 A:Accession: C83091
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-148 <STO>
 A:Cross-references: GB:AE004858; GB:AE004091; NID:g9950668; PIDN:AG07829.1; GSPDB:GN001;
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA4441

Query Match 3.7%; Score 7; DB 2; Length 148;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 89 ARLLPNA 95
 DB 26 ARLLPNA 32

RESULT 5
 AE2284
 Hypothetical protein alr3828 [imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp.
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
 C:Accession: AE2284
 A:Reference number: A81807; MUID:21595285; PMID:11759840
 A:Accession: AE2284
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-153 <KIR>
 A:Cross-references: GB:BA000019; PIDN:BA075527.1; PID:gl132962; GSPDB:GN00179
 A:Experimental source: strain PCC 7120
 C:Genetics:
 A:Gene: alr3828

Query Match 3.7%; Score 7; DB 2; Length 153;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 9 DATEIIP 15
 DB 66 DATEIIP 72

RESULT 6
 B37410
 H-2 class II histocompatibility antigen A-2 beta chain - spiny mouse (Mus saxicola) (frag
 C:Species: Mus saxicola
 C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 21-Jan-2000
 C:Accession: B37410
 A:Reference number: A37410; MUID:90316177; PMID:2369919
 A:Accession: B37410
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-183 <CAM>
 A:Cross-references: GB:J030158
 C:Superfamily: class II histocompatibility antigen; immunoglobulin homology
 F:105-170/Domain: immunoglobulin homology <IMM>
 Query Match 3.7%; Score 7; DB 2; Length 183;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 12 EIIEIIG 18

Db 157 EITPELG 163

RESULT 7

G81057

hypothetical protein NMB1656 [imported] - *Neisseria meningitidis* (strain MCS8 serogroup C)Species: *Neisseria meningitidis*

C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001

C/Accession: G81057

R/Tetcelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A. Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;

Science 287, 1809-1815, 2000

A/Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappaport, R.; V

A/Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MCS8.

A/Reference number: A81000; PMID:20175755; PMID:10710307

A/Accession: G81057

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-197 <TER>

A/Cross-references: GB:AE002516; GB:AE002098; NID:G7226905; PIDN:AAF42005.1; PID:G722690

A/Experimental source: serogroup B, strain MCS8

C/Genetics:

A/Gene: NMB1656

Query Match 3.7%; Score 7; DB 2; Length 197;

Best Local Similarity 100.0%; Pred. No. 22;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 KANQAEI 186

Db 122 KANQAEI 128

RESULT 8

A81819

hypothetical protein NMA1913 [imported] - *Neisseria meningitidis* (strain Z2491 serogroup C)Species: *Neisseria meningitidis*

C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001

C/Accession: A81819

R/Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel

Nature 404, 502-506, 2000

A/Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.

A/Reference number: A81775; PMID:20222556; PMID:10761919

A/Accession: A81819

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-197 <PAR>

A/Cross-references: GB:AL162757; GB:AL157959; NID:G7380371; PIDN:CAB85134.1; PID:G738054

A/Experimental source: serogroup A, strain Z2491

C/Genetics:

A/Gene: NMA1913

Query Match 3.7%; Score 7; DB 2; Length 197;

Best Local Similarity 100.0%; Pred. No. 22;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 KANQAEI 186

Db 122 KANQAEI 128

RESULT 9

A12930

hypothetical protein Atu3047 [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont) C/Species: *Agrobacterium tumefaciens*

C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002

C/Accession: A12930

R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L

Science 294, 2317-2323, 2001

A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E

A/Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.

A/Reference number: AB2577; PMID:11743193

A/Accession: A12930

A/Status: preliminary

A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E

A/Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.

A/Reference number: AB2577; PMID:11743193

A/Accession: A12930

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-296 <KUN>

A/Cross-references: GB:AE006889; PIDN:AA43863.1; PID:G17741408; GSPDB:GN00187

A/Experimental source: strain C58 (Dupont)

C/Genetics:

A/Gene: Atu3047

A/Map position: linear chromosome

C/Superfamily: oligopeptide permease protein oppB

Query Match 3.7%; Score 7; DB 2; Length 296;

Best Local Similarity 100.0%; Pred. No. 30;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 91 LLPNIG 97

Db 206 LLPNIG 212

RESULT 10

S77397

yabc protein homolog sl1144 - *Synechocystis* sp. (strain PCC 6803) C/Species: *Synechocystis* sp.

A/Variety: PCC 6803

C/Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000

C/Accession: S77397

R/Kaneko, T.; Sato, S.; Korani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;

DNA Res. 3, 109-136, 1996

A/Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*

S.

A/Reference number: S74322; PMID:97061201; PMID:8905231

A/Accession: S77397

A/Status: nucleic acid sequence not shown; translation not shown.

A/Molecule type: DNA

A/Residues: 1-305 <KAN>

A/Cross-references: EMBL:D90906; GB:AB001339; NID:G1652492; PIDN:BA17500.1; PID:G165257;

C/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C/Superfamily: *Escherichia coli* yabc protein

Query Match 3.7%; Score 7; DB 2; Length 305;

Best Local Similarity 100.0%; Pred. No. 31;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 174 PRARSAN 180

Db 292 PRARSAN 298

RESULT 11

E98351

oligopeptide transport system permease protein apbc AGR_L_3516 [imported] - *Agrobacterium*

C/Species: *Agrobacterium tumefaciens*

C/Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002

C/Accession: E98351

R/Goodner, B.; Hinkle, G.; Gatung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,

Science 294, 2323-2328, 2001

A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tume*

A/Reference number: A97359; PMID:11743194

A/Accession: E98351

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-311 <KUN>

A/Cross-references: GB:AB007870; PIDN:AAK90335.1; PID:G15160372; GSPDB:GN00170

C/Genetics:

A/Gene: AGR_L_3516

A/Map position: linear chromosome

C:Superfamily: oligopeptide permease protein oppb

Query Match 3.7%; Score 7; DB 2; Length 311;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 91 LLPNAIG 97
Db 221 LLPNAIG 227

RESULT 12

D75344 conserved hypothetical protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: D75344
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M.; Shen, M.; Vamathevan, J.D.; Lam, P.; McDonald, L.; Uitterback, T.; Zaleski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: D75344

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-312 <NHI>

A:Cross-references: GB:AE002026; GB:AE000513; NID:96459639; PIDN:AAF11419.1; PID:9645964

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR1866

A:Map position: 1

C:Superfamily: Escherichia coli yabc protein

Query Match 3.7%; Score 7; DB 2; Length 312;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 174 PRARSAK 180
Db 293 PRARSAK 299

RESULT 13

B82892 conserved hypothetical UUA17 [imported] - Ureaplasma urealyticum

C:Species: Ureaplasma urealyticum

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000

C:Accession: B82892

R:Glaser, J.L.; Lefkowitz, E.J.; Glaser, J.S.; Heiner, C.R.; Chen, E.Y.; Caswell, G.H.

submitted to Genbank, February 2000

A:Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a mit

A:Reference number: A82870

A:Accession: B82892

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-321 <GLA>

A:Cross-references: GB:AE002139; GB:AF222894; NID:96899405; PIDN:AAF30828.1; GSPDB:GN004

A:Experimental source: serovar 3; biovar 1

C:Genetics:

A:Gene: UUA17

A:Genetic code: SGC3

Query Match 3.7%; Score 7; DB 2; Length 321;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 135 KVRIVAS 141
Db 276 KVRIVAS 282

RESULT 14

S72647 hypothetical protein 1b - antracnose fungus (Colletotrichum gloeosporioides) retrotransi
C:Species: Glomerella cingulata, Colletotrichum gloeosporioides
C:Date: 24-Oct-1998 #sequence_revision 24-Oct-1998 #text_change 17-Mar-1999
C:Accession: S72647
R:He, C.; Nourse, J.P.; Kelemu, S.; Irwin, J.A.G.; Manners, J.M.
Mol. Gen. Genet. 252, 320-331, 1996
A:Title: CgT1: a non-LTR retrotransposon with restricted distribution in the fungal phyc
A:Reference number: S72619; MUID:96439839; PMID:8842152
A:Accession: S72647
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-328 <HBH>
A:Experimental source: biotype B, isolate UQ62
C:Genetics:
A:Mobile element: retrotransposon CgT1

Query Match 3.7%; Score 7; DB 2; Length 328;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 87 GPARLLP 93
Db 123 GPARLLP 129

RESULT 15

AG2834 SAM-dependent methyltransferase [imported] - Agrobacterium tumefaciens (strain C58, Dupont

C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002

C:Accession: AG2834

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Moo, L.

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell

; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; PMID:11743193

A:Accession: AG2834

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-341 <KUR>

A:Cross-references: GB:AE008688; PIDN:AAU43093.1; PID:917740563; GSPDB:GN00186

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: maw

A:Map position: circular chromosome

C:Superfamily: Escherichia coli yabc protein

Query Match 3.7%; Score 7; DB 2; Length 341;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 174 PRARSAK 180
Db 303 PRARSAK 309

Search completed: March 28, 2003, 14:23:57
Job time: 15.6933 secs

DR InterPro; IPR000359; Cys_knot.
 DR SMART; SMO0041; CT; 1.
 DR PROSITE; PS01185; CTCK_1; FALSE_NEG.
 DR PROSITE; PS01225; CTCK_2; FALSE_NEG.
 KW Signal; Glycoprotein.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 213 SCLEROSTIN.
 FT DOMAIN 82 172 CTCK.
 FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 213 AA; 24030 MW; 30DBD5CE73D5B2 CRC64;
 Query Match 100.0%; Score 190; DB 1; Length 213;
 Best Local Similarity 100.0%; Pred. No. 8.3e-195;
 Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 OGWOAFKNDATETIIPLEGYEPPEPELENNKTMNRAENGRRPHHPETKDVSEYSCREL 60
 DB 24 OGWOAFKNDATETIIPLEGYEPPEPELENNKTMNRAENGRRPHHPETKDVSEYSCREL 83
 QY 61 HFTRYVTDGPCRSAPVTELYVCSGCCGPARLLPNAIGRKWRRPSGDPFRCIPDRYRAOR 120
 DB 84 HFTRYVTDGPCRSAPVTELYVCSGCCGPARLLPNAIGRKWRRPSGDPFRCIPDRYRAOR 143
 QY 121 VOLLCPGGAPARARVRLVASCCKRLTRFNQSELKDFTEARPOKGRKRRPARSAK 180
 DB 144 VOLLCPGGAPARARVRLVASCCKRLTRFNQSELKDFTEARPOKGRKRRPARSAK 203
 QY 181 ANQAELENNY 190
 DB 204 ANQAELENNY 213
 RESULT 2
 SOST CERAE STANDARD; PRT; 213 AA.
 ID SOST CERAE
 AC 09BG78;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Sclerostin precursor.
 GN SOST.
 OS Cercopithecus aethiops (Green monkey) (Grivet).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Cercopithecus.
 NCBI_TaxID=9534;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21090529; PubMed=11179006;
 RA Brunkow M.E., Gardner J.C., Van Ness J., Paepker B.W., Kovacevich B.R.,
 RA Proll S., Skonier J.E., Zhao L., Sabo P.J., Fu Y.H., Alishch R.S.,
 RA Gillett L., Colbert T., Tacconi P., Galas D., Hamersma H.,
 RA Beighton P., Mulligan J.T.;
 RT "Bone dysplasia sclerosteosis results from loss of the SOST gene
 product, a novel cysteine knot-containing protein.";
 RL Am. J. Hum. Genet. 68:577-589(2001).
 CC -1- FUNCTION: Seems to play a role in bone homeostasis (By
 similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted (Potential).
 CC -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL; AF326742; AAK13457.1;
 DR InterPro; IPR000359; Cys_knot.
 DR SMART; SMO0041; CT; 1.

DR PROSITE; PS01185; CTCK_1; FALSE_NEG.
 DR PROSITE; PS01225; CTCK_2; FALSE_NEG.
 KW Signal; Glycoprotein.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 213 SCLEROSTIN.
 FT DOMAIN 82 172 CTCK.
 FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 213 AA; 23908 MW; 6DA7B5ED674728A CRC64;
 Query Match 67.4%; Score 128; DB 1; Length 213;
 Best Local Similarity 100.0%; Pred. No. 9.4e-129;
 Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 OGWOAFKNDATETIIPLEGYEPPEPELENNKTMNRAENGRRPHHPETKDVSEYSCREL 60
 DB 24 OGWOAFKNDATETIIPLEGYEPPEPELENNKTMNRAENGRRPHHPETKDVSEYSCREL 83
 QY 61 HFTRYVTDGPCRSAPVTELYVCSGCCGPARLLPNAIGRKWRRPSGDPFRCIPDRYRAOR 120
 DB 84 HFTRYVTDGPCRSAPVTELYVCSGCCGPARLLPNAIGRKWRRPSGDPFRCIPDRYRAOR 143
 QY 121 VOLLCPGG 128
 DB 144 VOLLCPGG 151
 RESULT 3
 SOST BOVIN STANDARD; PRT; 176 AA.
 ID SOST BOVIN
 AC 09BG79;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Sclerostin (Fragment).
 GN SOST.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21090529; PubMed=11179006;
 RA Brunkow M.E., Gardner J.C., Van Ness J., Paepker B.W., Kovacevich B.R.,
 RA Proll S., Skonier J.E., Zhao L., Sabo P.J., Fu Y.H., Alishch R.S.,
 RA Gillett L., Colbert T., Tacconi P., Galas D., Hamersma H.,
 RA Beighton P., Mulligan J.T.;
 RT "Bone dysplasia sclerosteosis results from loss of the SOST gene
 product, a novel cysteine knot-containing protein.";
 RL Am. J. Hum. Genet. 68:577-589(2001).
 CC -1- FUNCTION: Seems to play a role in bone homeostasis (By
 similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted (Potential).
 CC -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL; AF326738; AAK13453.1;
 DR InterPro; IPR000359; Cys_knot.
 DR SMART; SMO0041; CT; 1.
 DR PROSITE; PS01185; CTCK_1; FALSE_NEG.
 DR PROSITE; PS01225; CTCK_2; FALSE_NEG.
 KW Glycoprotein.
 FT NON TER 1 1
 FT DOMAIN 51 141 CTCK.
 FT CARBOHYD 22 22 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 144 144 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT NON_TER 176 176
 SQ SEQUENCE 176 AA; 19743 MW; 35FOCA61A425FADB CRC64;

Query Match 40.0%; Score 76; DB 1; Length 176;
 Best Local Similarity 100.0%; Pred. No. 2e-73;
 Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 SEYSCREHAFTRVYTDGPCRSAPKPVTELVCSGCGPARLLPNAIGKGMWRPSPDPFRCI 112
 DB 45 SEYSCREHAFTRVYTDGPCRSAPKPVTELVCSGCGPARLLPNAIGKGMWRPSPDPFRCI 104

QY 113 PDYRPAQRVQLCPGG 128
 DB 105 PDYRPAQRVQLCPGG 120

RESULT 4

SOST_RAT STANDARD; PRT; 213 AA.
 ID SOST_RAT
 AC 099P67;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Sclerostin precursor.
 GN SOST.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;

SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley;
 RX MEDLINE=21090529; PubMed=11179006;
 RA Brunkow M.E., Gardner J.C., Van Ness J., Paepers B.W., Kovacevich B.R.,
 RA Proll S., Skonier J.E., Zhao L., Sabo P.V., Fu Y.H., Alisch R.S.,
 RA Gillett L., Colbert T., Tacconi P., Galas D., Hamersma H.,
 RA Beighton P., Mulligan J.T.;
 RT "Bone dysplasia sclerosteosis results from loss of the SOST gene
 RT product, a novel cystine knot-containing protein.";
 RL Am. J. Hum. Genet. 68:577-589(2001).

- FUNCTION: Seems to play a role in bone homeostasis (By similarity).
 CC - SUBCELLULAR LOCATION: Secreted (Potential).
 CC - SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL; AF326741; AAK13456.1; -
 DR InterPro; IPR000359; Cys_knot.
 DR PROSITE; PS01185; CTCK_1; FALSE_NEG.
 DR PROSITE; PS01225; CTCK_2; FALSE_NEG.
 KW Signal; Glycoprotein.
 FT SIGNAL 1 28 POTENTIAL.
 FT CHAIN 29 213 SCLEROSTIN.
 FT DOMAIN 82 172 CTCK.
 FT CARBOHYD 53 53 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 175 175 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 213 AA; 23974 MW; 6C56C878C8D684B CRC64;

Query Match 17.4%; Score 33; DB 1; Length 213;
 Best Local Similarity 100.0%; Pred. No. 1.5e-27;
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 VTDGPCRSAPKPVTELVCSCGCGPARLLPNAIGR 98
 DB 89 VTDGPCRSAPKPVTELVCSCGCGPARLLPNAIGR 121

RESULT 5

SOST_MOUSE

ID SOST_MOUSE STANDARD; PRT; 211 AA.

AC 099P68; Q9D3L7;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Sclerostin precursor.
 GN SOST.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;

SEQUENCE FROM N.A.
 RC STRAIN=129/SVJ;
 RX MEDLINE=21090529; PubMed=11179006;
 RA Brunkow M.E., Gardner J.C., Van Ness J., Paepers B.W., Kovacevich B.R.,
 RA Proll S., Skonier J.E., Zhao L., Sabo P.V., Fu Y.H., Alisch R.S.,
 RA Gillett L., Colbert T., Tacconi P., Galas D., Hamersma H.,
 RA Beighton P., Mulligan J.T.;
 RT "bone dysplasia sclerosteosis results from loss of the SOST gene
 RT product, a novel cystine knot-containing protein.";
 RL Am. J. Hum. Genet. 68:577-589(2001).

SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Head;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schiomi L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barh G.,
 RA Blake J., Boffelli D., Bojunga N., Carinci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gueniche S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Saeki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kontsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).

- FUNCTION: Seems to play a role in bone homeostasis (By similarity).
 CC - SUBCELLULAR LOCATION: Secreted (Potential).
 CC - SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

EMBL; AF326740; AAK13455.1; -
 DR EMBL; AK017295; BAB30678.1; -
 DR EMBL; AF326737; AAK13452.1; -
 DR MGD; MGI:1921749; Sost.
 DR InterPro; IPR000359; Cys_knot.
 DR PROSITE; PS01185; CTCK_1; FALSE_NEG.
 DR PROSITE; PS01225; CTCK_2; FALSE_NEG.
 KW Signal; Glycoprotein.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 211 SCLEROSTIN.

FT DOMAIN 80 170 CTCK.
 FT CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 173 173 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CONFLICT 72 72 G -> D (IN REF. 2).
 SQ SEQUENCE 211 AA; 23443 MW; AEB094E358B34961 CRC64;
 Query Match 16.8%; Score 32; DB 1; Length 211;
 Best Local Similarity 100.0%; Pred. No. 1.8e-26;
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 67 TDGPCRSAPVTELVCSGCGPARLLPNAIGR 98
 Db 88 TDGPCRSAPVTELVCSGCGPARLLPNAIGR 119
 RESULT 6
 ELK1_HUMAN STANDARD; PRT; 428 AA.
 AC P19419; O75606; O9JUM4; O95058;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE ETS-domain protein ELK-1.
 GN ELK1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89203250; PubMed=2539641;
 RA Rao V.N., Huebner K., Isobe M., Ar-Rushdi A., Croce C.M.,
 RA Reddy E.S.P.;
 RT "Elk, tissue-specific, ets-related genes on chromosomes X and 14 near
 RT translocation breakpoints.";
 RL Science 244:66-70(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99013876; PubMed=9795224;
 RA Hardtman N., Mills F.C., Mitchell M.P., Meindl A., Max E.E.;
 RT "The human elk-1 gene family: the functional gene and two processed
 RT pseudogenes embedded in the Igh locus.";
 RL Gene 221:215-224(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Graftsm D.;
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RA Ayee D.N.T., Kovar H.;
 RT "Novel family members HuER71, ELFR, and ELKV among ETS-related genes
 RT coexpressed with EMS-FLI1 in Bwing tumor cell lines.";
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP DOMAINS.
 RX MEDLINE=92334979; PubMed=1630903;
 RA Janknecht R., Nordheim A.;
 RT "Elk-1 protein domains required for direct and SRF-assisted
 RT DNA-binding.";
 RL Nucleic Acids Res. 20:3317-3324(1992).
 CC -1- FUNCTION: STIMULATES TRANSCRIPTION. BINDS TO PURINE-RICH DNA
 CC SEQUENCES. CAN FORM A TERNARY COMPLEX WITH THE SERUM RESPONSE
 CC FACTOR AND THE ETS AND SRF MOTIFS OF THE FOS SERUM RESPONSE
 CC ELEMENT.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND 2/ELKV; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: LUNG AND TESTIS.
 CC -1- PTM: PHOSPHORYLATED BY THE THREE GROUPS OF MAP KINASE (P38, JNK,
 CC AND ERK).
 CC -1- SIMILARITY: BELONGS TO THE ETS FAMILY.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M25269; AAA52384.1; -;
 CC EMBL; AF080616; AAC82466.1; -;
 CC EMBL; AL009172; CAA15659.1; -;
 CC EMBL; AF000672; AAD00862.1; -;
 CC PIR; A41354; TVHUEK.
 CC HSSP; P28324; IBC8.
 CC TRASNFAIC; T00250; -;
 CC Genew; HGNC:3321; ELK1.
 CC MIM; 31040; -;
 CC InterPro; IPR000418; ETS.
 CC InterPro; IPR002341; HSF_ETs.
 CC Pfam; PF00178; Ets; 1.
 CC PRINTS; PR00454; ETSDOMAIN.
 CC SMART; SM00413; ETS; 1.
 CC PROSITE; PS00345; ETS_DOMAIN_1; 1.
 CC PROSITE; PS00346; ETS_DOMAIN_2; 1.
 CC PROSITE; PS50061; ETS_DOMAIN_3; 1.
 CC Transcription regulation; Activator; Nuclear protein; DNA-binding;
 CC Phosphorylation; Alternative splicing.
 CC DNA_BIND 5 86 ETS-DOMAIN.
 CC VARSPLIC 91 95 VAGCS -> SHCAP (IN ISOFORM 2).
 CC VARSPLIC 96 428 MISSING (IN ISOFORM 2).
 CC CONFLICT 183 183 S -> N (IN REF. 1).
 SQ SEQUENCE 428 AA; 44888 MW; 68F71F8ADBD9D38CA CRC64;
 Query Match 4.2%; Score 8; DB 1; Length 428;
 Best Local Similarity 100.0%; Pred. No. 1.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 166 PQGRRKPR 173
 Db 310 PQGRRKPR 317
 RESULT 7
 ELK1_MOUSE STANDARD; PRT; 429 AA.
 AC P41969;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE ETS-domain protein ELK-1.
 GN ELK1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=C57BL/6; TISSUE=Embryo;
 RX MEDLINE=97017146; PubMed=8663747;
 RA Grevin D., Ung S., Demuez F., Denem M., Quatmanns B., Begue A.,
 RA Stenelin D., Martin P.;
 RT "Structure and organization of the mouse elk1 gene.";
 RL Gene 174:185-188(1996).
 RN [2]
 RP SEQUENCE OF 5-224 FROM N.A.
 RX TISSUE=Embryo;
 RX MEDLINE=95047310; PubMed=7958635;
 RA Giovane A., Pintzas A., Maira S.-M., Sobieszczuk P., Waeslyk B.;
 RT "Net, a new ets transcription factor that is activated by Ras.";
 RL Gene Dev. 8:1507-1513(1994).
 CC -1- FUNCTION: STIMULATES TRANSCRIPTION. BINDS TO PURINE-RICH DNA
 CC SEQUENCES. CAN FORM A TERNARY COMPLEX WITH THE SERUM RESPONSE
 CC FACTOR AND THE ETS AND SRF MOTIFS OF THE FOS SERUM RESPONSE
 CC ELEMENT.

```

CC -|- SUBCELLULAR LOCATION: Nuclear.
CC -|- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN THE BRAIN, AND TO A
CC LESSER EXTENT IN THE HEART, LIVER AND MUSCLE.
CC -|- SIMILARITY: BELONGS TO THE ETS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X87257; CAA60715.1; -
DR EMBL; Z36939; CAA85391.1; -
DR HSSP; P28324; IBC8.
DR TRANSFAC; T05013; -.
DR MGD; MGI:101833; Elk1.
DR InterPro; IPR000418; Ets.
DR InterPro; IPR002341; HSF_ETs.
DR Pfam; PF00178; Ets; 1.
DR PRINTS; PR00454; ETSDOMAIN.
DR SMART; SM00413; ETS; 1.
DR PROSITE; PS00345; ETS_DOMAIN_1; 1.
DR PROSITE; PS00346; ETS_DOMAIN_2; 1.
DR PROSITE; PS50061; ETS_DOMAIN_3; 1.
KW Transcription regulation; Activator; Nuclear protein; DNA-binding;
KW Phosphorylation.
FT DNA_BIND 5 86 ETS-DOMAIN.
FT CONFLICT 133 133 P -> T (IN REF. 2).
SQ SEQUENCE 429 AA; 45243 MW; B61B5B97731054F CRC64;

Query Match
Best Local Similarity 100.0%; Score 8; DB 1; Length 429;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 166 POKGRKR 173
Db 311 POKGRKR 318

RESULT 8
RL22_LEPIN
ID RL22_LEPIN STANDARD; PRT; 110 AA.
AC Q9XJ31;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE 50S ribosomal protein L22.
GN RPLV.
OS Leptospira interrogans.
OC Bacteria; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxId=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Setovar Lai;
RX MEDLINE=20088835; PubMed=10620683;
RA Zuercher R.L., Hartskeerl R.A., van de Kemp H., Bal A.3.;
RT "Characterization of the Leptospira interrogans S10-spcc-alpha
RT operon."
RL FEMS Microbiol. Lett. 182:303-308(2000).
CC -|- FUNCTION: THIS PROTEIN BINDS SPECIFICALLY TO 23S rRNA; ITS BINDING
CC IS STIMULATED BY OTHER RIBOSOMAL PROTEINS, E.G., L4, L17, AND L20.
CC IT IS IMPORTANT DURING THE EARLY STAGES OF 50S RECONSTITUTION
CC (BY SIMILARITY).
CC -|- SIMILARITY: BELONGS TO THE L22P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF15283; AAD40588.1; -
DR HSSP; P48286; IBE.
DR InterPro; IPR001063; Ribosomal_L22.
DR Pfam; PF00237; Ribosomal_L22; 1.
DR ProDom; PD001032; Ribosomal_L22; 1.
DR TIGRFAMs; TIGR01044; rplV_bact; 1.
DR PROSITE; PS00464; RIBOSOMAL_L22; 1.
KW Ribosomal protein; rRNA-binding
SQ SEQUENCE 110 AA; 12566 MW; B9181B85E9756CAA CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 1; Length 110;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 134 RYRLVA 140
Db 15 RYRLVA 21

RESULT 9
MRW_SYNY3
ID MRW_SYNY3 STANDARD; PRT; 305 AA.
AC P73460;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE S-adenosyl-methyltransferase mrw (EC 2.1.1.-).
GN MRW OR SLI114.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxId=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirosewa M., Sugitara M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naito K., Okumura S.,
RA Shimpō S., Takeuchi C., Wada T., Watanabe A., Yanada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions."
RL DNA Res. 3:109-136(1996).
CC -|- FUNCTION: Exhibits a S-adenosyl-dependent methyltransferase
CC activity (By similarity).
CC -|- SIMILARITY: BELONGS TO THE MRW FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D90906; BAA17500.1; -
DR InterPro; IPR002903; Bac_Metnfrse.
DR Pfam; PF01795; Methyltransf_5; 1.
DR ProDom; PD004685; Bac_Metnfrse; 1.
DR TIGRFAMs; TIGR00006; UPF0117; 1.
KW Transferase; Methyltransferase; Complete proteome.
SQ SEQUENCE 305 AA; 34386 MW; 758F9F567247129F CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 1; Length 305;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 174 PRASAK 180
Db 292 PRASAK 298

```

```

RESULT 10
MRW DEIRA          STANDARD;      PRT;      312 AA.
ID MRW DEIRA
AC Q9RT99;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE S-adenosyl-methyltransferase mraw (EC 2.1.1.-).
GN MRAW OR DR1866.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RI;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Bisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1."
RL Science 286:1571-1577(1999).
CC - FUNCTION: Exhibits a S-adenosyl-dependent methyltransferase
CC activity (By similarity).
CC - SIMILARITY: BELONGS TO THE MRAW FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE002026; AAF1419.1; -.
DR TIGR; DR1866; -.
DR InterPro; IPR002903; Bac_Metnfrase.
DR Pfam; PF01795; Methyltransf_5; 1.
DR ProDom; PD004648; Bac_Metnfrase; 1.
DR TIGRFAMs; TIGR00006; UPP0117; 1.
KW Transferase; Methyltransferase; Complete proteome.
SQ SEQUENCE 312 AA; 33443 MW; B51835F2273A9463 CRC64;

Query Match          3.7%; Score 7; DB 1; Length 312;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 174 PRARSAX 180
Db 293 PRARSAX 299

RESULT 11
MRW AGRT5          STANDARD;      PRT;      341 AA.
ID MRW AGRT5
AC P58745;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE S-adenosyl-methyltransferase mraw (EC 2.1.1.-).
GN MRAW OR ATU2102 OR AGR C 3815.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boyce D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-O., McClelland E., Palmeri A., Gordon D.,
RA Raymond C., Rouse G., Saenphitumachak C., Wu Z., Romero P., Perry M.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley P., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Neeter E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58."
RL Science 294:2317-2323 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194; Miller N., Blanchard M.,
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houniel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughy D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Garson J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58."
RL Science 294:2323-2328(2001).
CC - FUNCTION: Exhibits a S-adenosyl-dependent methyltransferase
CC activity (By similarity).
CC - SIMILARITY: BELONGS TO THE MRAW FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE009160; AAL43093.1; -.
DR EMBL; AE008126; AAK87852.1; ALT INIT.
DR InterPro; IPR002903; Bac_Metnfrase.
DR Pfam; PF01795; Methyltransf_5; 1.
DR TIGRFAMs; TIGR00006; UPP0117; 1.
KW Transferase; Methyltransferase; Complete proteome.
SQ SEQUENCE 341 AA; 36369 MW; 5EED5F0731B8C208 CRC64;

Query Match          3.7%; Score 7; DB 1; Length 341;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 174 PRARSAX 180
Db 303 PRARSAX 309

RESULT 12
MRW RHIME          STANDARD;      PRT;      341 AA.
ID MRW RHIME
AC Q92N14;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE S-adenosyl-methyltransferase mraw (EC 2.1.1.-).
GN MRAW OR R02184 OR SMC01858.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,

```

RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
 RA Godtard T., Goffeau A., Kahn D., Käss B., Lelaure V., Masny D.,
 RA Pohl T., Poretelle D., Puehler A., Purnelle B., Ramsperger U.,
 RA Renard C., Thebaud P., Vandenbol M., Weidner S., Galibert F.,
 RT "Analysis of the chromosome sequence of the legume symbiont
 RT *Sinorhizobium meliloti* strain 1021.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
 CC -I- FUNCTION: Exhibits a S-adenosyl-dependent methyltransferase
 CC activity (By similarity).
 CC -I- SIMILARITY: BELONGS TO THE MRAM FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AL591789; CAC46763.1; -;
 DR InterPro; IPR002903; Bac_Metrifree.
 DR Pfam; PF01795; Methyltransf_5; 1.
 DR ProDom; PD004685; Bac_Metrifree; 1.
 DR TIGRFAMs; TIGR00006; UpP0117; 1.
 DR Trnstrfase; Methyltransferase; Complete proteome.
 KW Trnstrfase; Methyltransferase; Complete proteome.
 SQ SEQUENCE 341 AA; 36021 MW; D84227F3D9B7938 CRC64;
 Query Match 3.7%; Score 7; DB 1; Length 341;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 174 PRARSAK 180
 DB 303 PRARSAK 309
 RESULT 13
 MRAM BRUWE STANDARD; PRT; 346 AA.
 AC Q8Y174;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE S-adenosyl-methyltransferase mraw (EC 2.1.1.-).
 GN MRAW OR BMEI0571.
 OS Brucella melitensis.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 CC Brucellaceae; Brucella.
 CC NCB1_Taxid=29459;
 CC (1)
 RN SEQUENCE FROM N.A.
 RP STRAIN=16M / ATCC 23456 / Biotype 1;
 RX MEDLINE=20020109; PubMed=11756688;
 RA Delvecchio V.G., Kapratral V., Redkar R.J., Patra G., Mujer C., Los T.,
 RA Ivanova N., Andersen T., Bhattacharya A., Lykidis A., Resnik G.,
 RA Jablonkai L., Larsen N., D Souza M., Bernal A., Mazur M., Goldsman E.,
 RA Selkov E., Elser P.H., Hagius S., O'Callaghan D., Lelsson J.-J.,
 RA Haselkorn R., Kyriades N., Overbeek R.;
 RT "The genome sequence of the facultative intracellular pathogen
 RT *Brucella melitensis*.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
 CC -I- FUNCTION: Exhibits a S-adenosyl-dependent methyltransferase
 CC activity (By similarity).
 CC -I- SIMILARITY: BELONGS TO THE MRAM FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AE009499; AL51752.1; -;
 DR InterPro; IPR002903; Bac_Metrifree.
 DR Pfam; PF01795; Methyltransf_5; 1.
 DR ProDom; PD004685; Bac_Metrifree; 1.
 DR TIGRFAMs; TIGR00006; UpP0117; 1.
 DR Trnstrfase; Methyltransferase; Complete proteome.
 KW Trnstrfase; Methyltransferase; Complete proteome.
 SQ SEQUENCE 346 AA; 37501 MW; 23F15D1F05D7BB6 CRC64;
 Query Match 3.7%; Score 7; DB 1; Length 346;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 174 PRARSAK 180
 DB 307 PRARSAK 313
 RESULT 14
 NUSM DIDMA STANDARD; PRT; 602 AA.
 ID NUSM DIDMA
 AC P41309; P92474;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE NADH-ubiquinone oxidoreductase chain 5 (EC 1.6.5.3).
 GN MTRND5 OR ND5.
 OS Didelphis marsupialis virginiana (North American opossum).
 CC Mitochondrion.
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Metatheria; Didelphimorphia; Didelphidae; Didelphis.
 CC NCB1_Taxid=9267;
 CC (1)
 RN SEQUENCE FROM N.A.
 RP TISSUE=Liver.
 RC MEDLINE=94333786; PubMed=8056314;
 RA Janke A., Feldmaier-Fuchs G., Thomas K., von Haeseler A., Pabo S.;
 RT "The marsupial mitochondrial genome and the evolution of placental
 RT mammals.";
 RL Genetics 137:243-256(1994).
 CC -I- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z29573; CA82687.1; -;
 DR PIR; S42820; S42820.
 DR InterPro; IPR003916; NADH_oxred5.
 DR InterPro; IPR001750; Oxidored_q1.
 DR InterPro; IPR001516; Oxidored_q1_N.
 DR Pfam; PF00361; Oxidored_q1; 1.
 DR Pfam; PF00662; Oxidored_q1_N; 1.
 DR PRINTS; PR01434; NADH_OXNASE5.
 KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion.
 SQ SEQUENCE 602 AA; 68001 MW; 1020P172592669HE CRC64;
 Query Match 3.7%; Score 7; DB 1; Length 602;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 27 LENKTM 33
 DB 267 LENKTM 273
 RESULT 15
 VGLG SYNV STANDARD; PRT; 632 AA.
 ID VGLG SYNV
 AC P27277;

DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-JUL-1993 (Rel. 26, Last annotation update)
DE Spike glycoprotein precursor.
GN G.
OS Sonchus yellow net virus (SYNV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Nucleorhabdovirus.
OX NCBI_TaxID=11307;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC PV-263;
RX MEDLINE=92024089; PubMed=1926779;
RA Goldberg K.B., Modrell B., Hillman B.I., Heaton L.A., Choi T.J.,
RA Jackson A.O.;
RT "Structure of the glycoprotein gene of sonchus yellow net virus, a
RT plant rhabdovirus.";
RL Virology 185:32-38(1991).
CC -1- FUNCTION: THIS PROTEIN FORMS SPIKES ON THE SURFACE OF THE VIRION.
CC IT IS RESPONSIBLE BOTH FOR THE BINDING OF THE VIRUS TO SUSCEPTIBLE
CC HOST CELLS AND FOR INDUCING THE UPTAKE OF THE VIRUS BY THE CELL.
CC THE INTERACTION BETWEEN THE INTERNAL COMPONENTS OF THE VIRION
CC AND THE PORTION OF THE GLYCOPROTEIN EXPOSED ON THE CYTOPLASMIC
CC FACE OF THE PLASMA MEMBRANE PROBABLY DIRECTS ENVELOPMENT AND
CC VIRUS BUDDING.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L32603; AAA50384.1; -.
DR EMBL; M73626; AAA47898.1; -.
DR PIR; A40776; VGVNSY.
KW Transmembrane; Envelope protein; Glycoprotein; Signal.
FT STGNL 1 17 POTENTIAL.
FT CHAIN 18 632 SPIKE GLYCOPROTEIN.
FT TRANSMEM 560 578 POTENTIAL.
FT CARBOHYD 93 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 385 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 501 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 512 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 541 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 632 AA; 71093 MW; 05541E78BEA07927 CRC64;
Query Match 3.7%; Score 7; DB 1; Length 632;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 135 KVRIVAS 141
|||||

DB 579 KVRIVAS 585

Search completed: March 28, 2003, 14:22:21
Job time : 9.10667 secs

GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:19:52 ; Search time 26.8533 Seconds
(without alignments)
1457.881 Million cell updates/sec

Title: US-09-867-274-2

Perfect score: 190
Sequence: 1 QGMOAFKNDATEIPELGEY.....KRPBRASAKANOAELENAY 190

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size: 0

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	9	4.7	206	4 Q96HJ7	Q96HJ7 homo sapien
2	9	4.7	206	11 Q96HJ7	Q96HJ7 mus musculu
3	9	4.7	206	11 Q96HJ7	Q96HJ7 oryza sativ
4	8	4.2	276	13 Q9YGD6	Q9YGD6 oncorhynch
5	8	4.2	276	13 Q9YGD6	Q9YGD6 oncorhynch
6	8	4.2	276	13 Q9YGD6	Q9YGD6 oncorhynch
7	8	4.2	276	13 Q9YGD6	Q9YGD6 oncorhynch
8	8	4.2	276	13 Q9YGD6	Q9YGD6 oncorhynch
9	8	4.2	276	13 Q9YGD6	Q9YGD6 oncorhynch
10	8	4.2	276	13 Q9YGD6	Q9YGD6 oncorhynch
11	8	4.2	276	13 Q9YGD6	Q9YGD6 oncorhynch
12	8	4.2	276	13 Q9YGD6	Q9YGD6 oncorhynch
13	8	4.2	276	13 Q9YGD6	Q9YGD6 oncorhynch
14	8	4.2	276	13 Q9YGD6	Q9YGD6 oncorhynch
15	8	4.2	276	13 Q9YGD6	Q9YGD6 oncorhynch
16	8	4.2	276	13 Q9YGD6	Q9YGD6 oncorhynch

17	7	3.7	197	16 Q9JTA0	Q9JTA0 neisseria m
18	7	3.7	214	10 Q8RUS7	Q8RUS7 oryza sativ
19	7	3.7	251	10 Q9LINO	Q9LINO streptococ
20	7	3.7	282	10 Q9XFI9	Q9XFI9 arabidopsi
21	7	3.7	311	16 Q8UBG9	Q8UBG9 agrobacteri
22	7	3.7	321	16 Q9FQ72	Q9FQ72 ureaplasma
23	7	3.7	367	11 Q9IYB3	Q9IYB3 mus musculu
24	7	3.7	398	11 Q00388	Q00388 methanococ
25	7	3.7	425	5 Q19348	Q19348 caenorhabdi
26	7	3.7	562	10 Q81837	Q81837 arabidopsi
27	7	3.7	567	16 Q9A929	Q9A929 caulobacter
28	7	3.7	604	5 Q44003	Q44003 toxoplasma
29	7	3.7	632	4 Q8WXA1	Q8WXA1 homo sapien
30	7	3.7	632	4 Q96K34	Q96K34 homo sapien
31	7	3.7	632	16 Q92D98	Q92D98 listeria in
32	7	3.7	638	4 Q9H7Y7	Q9H7Y7 homo sapien
33	7	3.7	648	4 Q9HAI2	Q9HAI2 homo sapien
34	7	3.7	667	4 Q9Y2N7	Q9Y2N7 homo sapien
35	7	3.7	681	16 Q9PGB8	Q9PGB8 xylella fas
36	7	3.7	739	5 Q9BSC4	Q9BSC4 homo sapien
37	7	3.7	739	5 Q9W4Z3	Q9W4Z3 drosophila
38	7	3.7	782	5 Q9NEH9	Q9NEH9 drosophila
39	7	3.7	802	5 Q96398	Q96398 schistosoma
40	7	3.7	815	4 Q43273	Q43273 homo sapien
41	7	3.7	917	11 Q54892	Q54892 rattus norv
42	7	3.7	962	17 Q8TQC1	Q8TQC1 methanosarc
43	7	3.7	974	5 Q20143	Q20143 caenorhabdi
44	7	3.7	980	4 Q9NS55	Q9NS55 homo sapien
45	7	3.7	1008	16 Q92UM7	Q92UM7 rhizobium m

ALIGNMENTS

RESULT 1

Q96HJ7 PRELIMINARY; PRT; 206 AA.

AC Q96HJ7;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Similar to RIKEN cDNA 061006G05 gene (CDA019).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BONE MARROW;
RA Strauberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Liu F., Xu X.R., Qian B.Z., Xiao H., Chen Z., Han Z.;
RT "A novel gene expressed in human pheochromocytoma."
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC008484; AA08484.1; -
DR EMBL; AF361494; AAL57219.1; -
DR InterPro; IPR000359; Cys_knot.
DR PROSITE; PS01225; CTCK_2; 1.
SQ SEQUENCE 206 AA; 23306 MW; 9FBJCC41EAB53834 CRC64;

Query Match 4.7%; Score 9; DB 4; Length 206;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 AFKNDATEI 13
Db 23 AFKNDATEI 31

RESULT 2
Q9CON4

```

ID 09CON4 PRELIMINARY; PRT; 206 AA.
AC 09CON4;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE 0610006G05Rik protein (RIKEN CDNA 0610006G05 gene).
GN 0610006G05Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CS7BL/6J; TISSUE=PANCREAS AND KIDNEY;
RX MEDLINE=21085660; Pubmed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Akakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamamaki I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Knehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Wastio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldi M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Guelincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli U., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Wetz C., Whitaker C., Wilmig L.,
RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK007967; BAB25378.1; -
DR EMBL; AK002240; BAB21957.1; -
DR EMBL; AK002396; BAB22068.1; -
DR EMBL; AK007693; BAB25333.1; -
DR EMBL; BC021458; AAB21458.1; -
DR WGD; MGI:1913292; 0610006G05Rik.
DR InterPro; IPR000359; Cys_knoc.
DR PROSITE; PS01225; CTCK_2; 1.
SQ SEQUENCE 206 AA; 23174 MW; 70D24819EB06CBC9 CRC64;

Query Match 4.7%; Score 9; DB 11; Length 206;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 AFKDATEI 13
DB 23 AFKDATEI 31

RESULT 3
09FY03 PRELIMINARY; PRT; 198 AA.
AC 09FY03;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Hypothetical protein (P0011G08.32 protein).
GN P0011G08.32.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Eubartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nippobare (GA3) genomic DNA, chromosome 1, PAC
RT clone:P0433F09."
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nippobare (GA3) genomic DNA, chromosome 1, PAC
RT clone:P0011G08."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP002539; BAB08182.1; -
DR EMBL; AP003225; BAB64660.1; -
DR InterPro; IPR000886; ER target.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 198 AA; 21843 MW; 11DD190FE4B72FAE CRC64;

Query Match 4.2%; Score 8; DB 10; Length 198;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 126 PGSEAPRA 133
DB 83 PGSEAPRA 90

RESULT 4
09YGD6 PRELIMINARY; PRT; 276 AA.
AC 09YGD6;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Carboxyl reductase/20beta-hydroxysteroid dehydrogenase B.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Procaranthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99183507; Pubmed=10082666;
RA Guan G., Tanaka M., Todo T., Young G., Yoshikuni M., Nagahama Y.;
RT "Cloning and expression of two carboxyl reductase-like 20beta-
RT hydroxysteroid dehydrogenase cDNAs in ovarian follicles of rainbow
RT trout."
RL Biochem. Biophys. Res. Commun. 255:123-128(1999).
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
DR EMBL; AF100930; AAD20217.1; -
DR InterPro; IPR002198; ADH short.
DR Pfam; PF00106; adh short; 1.
DR PRINTS; PR00080; SDRFAMILY.
KW Oxidoreductase.
SQ SEQUENCE 276 AA; 30111 MW; 135604A38D2940CD CRC64;

Query Match 4.2%; Score 8; DB 13; Length 276;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 AFKDATE 12
DB 93 AFKDATE 100

RESULT 5
09PT36 PRELIMINARY; PRT; 276 AA.
AC 09PT36;

```


DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Carboxyl reductase/20beta-hydroxysteroid dehydrogenase A.
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 CC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 NCBI_TaxID=8022;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99185307; PubMed=10082666;
 RA Guan G., Tanaka M., Todo T., Young G., Yoshikuni M., Nagahama Y.;
 RT "Cloning and expression of two carboxyl reductase-like 20beta-
 RT hydroxysteroid dehydrogenase cDNAs in ovarian follicles of rainbow
 RT trout";
 RL Biochem. Biophys. Res. Commun. 255:123-128(1999).
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 CC (SDR) FAMILY.
 CC EMBL; AF100931; AAD20218.1; -.
 DR HSSP; P50162; IAE1.
 DR InterPro: IPR002198; ADH short.
 DR Pfam; PF00106; adh_short; 1.
 DR PRINTS; PR00080; SDRFAMILY.
 KW Oxidoreductase.
 SQ SEQUENCE 276 AA; 30139 MW; 135F1711D30086D CRC64;

Query Match 4.2%; Score 8; DB 13; Length 276;
 Best Local Similarity 100.0%; Pred. No. 5.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 AFKNDATE 12
 DB 93 AFKNDATE 100

RESULT 6
 Q9PT37 PRELIMINARY; PRT; 276 AA.
 ID Q9PT37
 AC Q9PT37;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Carboxyl reductase/20beta-hydroxysteroid dehydrogenase B.
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 CC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 NCBI_TaxID=8022;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99185307; PubMed=10082666;
 RA Guan G., Tanaka M., Todo T., Young G., Yoshikuni M., Nagahama Y.;
 RT "Cloning and expression of two carboxyl reductase-like 20beta-
 RT hydroxysteroid dehydrogenase cDNAs in ovarian follicles of rainbow
 RT trout";
 RL Biochem. Biophys. Res. Commun. 255:123-128(1999).
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 CC (SDR) FAMILY.
 CC EMBL; AF100932; AAD20991.1; -.
 DR InterPro: IPR002198; ADH short.
 DR Pfam; PF00106; adh_short; 1.
 DR PRINTS; PR00080; SDRFAMILY.
 KW Oxidoreductase.
 SQ SEQUENCE 276 AA; 30127 MW; 1350F4211D34225D CRC64;

Query Match 4.2%; Score 8; DB 13; Length 276;
 Best Local Similarity 100.0%; Pred. No. 5.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 AFKNDATE 12
 DB 93 AFKNDATE 100

RESULT 7
 Q9PT38 PRELIMINARY; PRT; 276 AA.
 ID Q9PT38
 AC Q9PT38;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Carboxyl reductase/20beta-hydroxysteroid dehydrogenase A.
 OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 CC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 NCBI_TaxID=8022;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99185307; PubMed=10082666;
 RA Guan G., Tanaka M., Todo T., Young G., Yoshikuni M., Nagahama Y.;
 RT "Cloning and expression of two carboxyl reductase-like 20beta-
 RT hydroxysteroid dehydrogenase cDNAs in ovarian follicles of rainbow
 RT trout";
 RL Biochem. Biophys. Res. Commun. 255:123-128(1999).
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 CC (SDR) FAMILY.
 CC EMBL; AF100933; AAD20992.1; -.
 DR HSSP; P50162; IAE1.
 DR InterPro: IPR002198; ADH short.
 DR Pfam; PF00106; adh_short; 1.
 DR PRINTS; PR00080; SDRFAMILY.
 KW Oxidoreductase.
 SQ SEQUENCE 276 AA; 30239 MW; 134B52054C751C28 CRC64;

Query Match 4.2%; Score 8; DB 13; Length 276;
 Best Local Similarity 100.0%; Pred. No. 5.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 AFKNDATE 12
 DB 93 AFKNDATE 100

RESULT 8
 Q969X8 PRELIMINARY; PRT; 428 AA.
 ID Q969X8
 AC Q969X8;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE ELK1 protein.
 GN ELK1.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=HIPPOCAMPUS;
 RX MEDLINE=99246057; PubMed=10231026;
 RA Yanahuchi T., Toko M., Suga M., Hatakeyama T., Isebe M.;
 RT "Structural organization of the human ELK1 gene and its processed
 RT pseudogene ELK2 genes";
 RL DNA Res. 6:21-27(1999).
 DR EMBL; AB016194; BAA36617.1; -.
 DR EMBL; AB016193; BAA36616.1; -.
 DR InterPro: IPR000418; Ets.
 DR InterPro: IPR002341; HSF_ETs.
 DR Pfam; PF00178; Ets; 1.
 DR PROSITE; PS00345; ETS DOMAIN 1; UNKNOWN 1.
 DR PROSITE; PS00346; ETS DOMAIN 2; UNKNOWN 1.
 DR PROSITE; PS50061; ETS_DOMAIN_3; 1.
 SQ SEQUENCE 428 AA; 44915 MW; 58FBFBAD9895A83 CRC64;

Query Match 4.2%; Score 8; DB 4; Length 428;
 Best Local Similarity 100.0%; Pred. No. 8.4;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 POKGRKPR 173
 |||||
 DB 310 POKGRKPR 317

RESULT 9

Q9FH55 PRELIMINARY; PRT; 89 AA.
 ID Q9FH55
 AC Q9FH55;
 DT 01-MAR-2001 (TRENBLrel. 16, Created)
 DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
 DE Genomic DNA, chromosome 5, TAC clone:K1L20.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
 NC NCB1_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RX MEDLINE=20181125; Pubmed=10718197;
 RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
 RA Tsubata S.,
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
 RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
 RT clones."
 RL DNA Res. 7:31-63 (2000).
 DR EMBL: AB022211; BAB10713.1;
 SQ SEQUENCE 89 AA; 10345 MW; DP12344C836FAE0A CRC64;

Query Match 3.7%; Score 7; DB 10; Length 89;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 90 RLPLNPI 96
 |||||
 DB 18 RLPLNPI 24

RESULT 10

Q9GTU4 PRELIMINARY; PRT; 109 AA.
 ID Q9GTU4
 AC Q9GTU4;
 DT 01-MAR-2001 (TRENBLrel. 16, Created)
 DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
 DE Zinc metalloproteinase 3 ME3 (Fragment).
 OS Ancylostoma caninum (Dog hookworm).
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
 OC Ancylostomatoidae; Ancylostomatidae; Ancylostomatinae; Ancylostoma.
 NC NCB1_TaxID=29170;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Jones B.F., Hotez P.J.,
 RT "Cloning and characterization of a zinc-metalloprotease secreted by
 RT the invasive stages of Ancylostoma caninum."
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF273706; AAG29106.1;
 DR HSSP; P08473; 1DMT.
 DR InterPro; IPR000718; Peptidase_M13.
 DR Pfam; PF01431; Peptidase_M13; 1.
 FT NON TER 1
 SQ SEQUENCE 109 AA; 12300 MW; 7C3C54FD724CAD7B CRC64;

Query Match 3.7%; Score 7; DB 5; Length 109;
 Best Local Similarity 100.0%; Pred. No. 30;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 108 DRCIPD 114
 |||||
 DB 86 DRCIPD 92

RESULT 11

Q910C0 PRELIMINARY; PRT; 130 AA.
 ID Q910C0
 AC Q910C0;
 DT 01-MAR-2001 (TRENBLrel. 16, Created)
 DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
 DT 01-OCT-2001 (TRENBLrel. 18, Last annotation update)
 DE Hypothetical protein PA2722.
 GN PA2722.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 NC NCB1_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; Pubmed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brady L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reiter J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.,
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RT opportunistic pathogen."
 RL Nature 406:959-964 (2000).
 DR EMBL: AE004700; AAG06110.1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 130 AA; 14641 MW; ECEBDCD7F247835E CRC64;

Query Match 3.7%; Score 7; DB 16; Length 130;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 135 KYRLVAS 141
 |||||
 DB 12 KYRLVAS 18

RESULT 12

Q9HVX4 PRELIMINARY; PRT; 148 AA.
 ID Q9HVX4
 AC Q9HVX4;
 DT 01-MAR-2001 (TRENBLrel. 16, Created)
 DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
 DT 01-OCT-2001 (TRENBLrel. 18, Last annotation update)
 DE Hypothetical protein PA4441.
 GN PA4441.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 NC NCB1_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; Pubmed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brady L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reiter J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.,
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RT opportunistic pathogen."
 RL Nature 406:959-964 (2000).
 DR EMBL: AE004858; AAG07829.1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 148 AA; 16418 MW; 4DD9E4B217EDD2B8 CRC64;

Query Match 3.7%; Score 7; DB 16; Length 148;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 89 ARLLPNA 95
 DB 26 ARLLPNA 32

RESULT 13
 ID Q8YQJ5 PRELIMINARY; PRT; 153 AA.

AC Q8YQJ5; 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DE Hypothetical protein A13828.

OS Anabaena sp. (strain PCC 7120).
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 NCBI_TaxId=103690;

RP SEQUENCE FROM N.A.
 RX MEDLINE=21595285; PubMed=11759840; Kuritz T., Sasamoto S.,
 RA Kaneko T., Nakamura Y., Wolk C.P., Kawashima K., Kimura T.,
 RA Watanabe A., Iriyuchi M., Ishikawa A., Matsuno A., Muraki A.,
 RA Kishida Y., Kohara M., Matsumoto M., Takazawa M., Yamada M.,
 RA Nakazaki N., Shimo S., Sugimoto M.,
 RA Yasuda M., Tabata S.;

RT "Complete genomic sequence of the filamentous nitrogen-fixing
 cyanobacterium Anabaena sp. strain PCC 7120.";
 RL DNA Res. 8:205-213(2001).

DR EMBL: AP003594; BAB75527.1;
 DR InterPro: IPR003728; DUF150;

KW Pfam: PF02576; DUF150; 1.
 KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 153 AA; 17259 MW; EBD82DIFED43FC6 CRC64;

Query Match 3.7%; Score 7; DB 16; Length 153;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 DATEIIP 15
 DB 66 DATEIIP 72

RESULT 14
 ID Q31243 PRELIMINARY; PRT; 183 AA.

AC Q31243; 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DE MHC class II A-beta 2 (Fragment).

OS Mus saxicola (Spiry mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_TaxId=10094;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90316177; PubMed=2369919;
 RA Cam P., Jovin-Marche E., Leguenn C., Marche P.N.;

RT "Structure of class II genes in wild mouse Mus saxicola: Functional
 and evolutionary implications.";
 RL Eur. J. Immunol. 20:1337-1343(1990).

DR EMBL: M30157; AAA39585.1;
 DR EMBL: M30156; AAA39585.1; JOINED.
 DR HSSP: P13760; 2SER.
 DR InterPro: IPR003597; IG_C1.
 DR InterPro: IPR003066; IG_MHC.
 DR InterPro: IPR000353; MHC_II_beta.

DR Pfam: PF00047; Ig_1.
 DR Pfam: PF00969; MHC_II_beta.1.
 DR ProDom: PD000328; MHC_II_beta.1.
 DR SMART: SM00407; IGc1_1.
 DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
 KW Glycoprotein; MHC II; Transmembrane.

FT NON TER 1
 SQ SEQUENCE 183 AA; 20906 MW; 79471BF03050AAC9 CRC64;

Query Match 3.7%; Score 7; DB 7; Length 183;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 EIPPLG 18
 DB 157 EIPPLG 163

RESULT 15
 ID Q9SDR7 PRELIMINARY; PRT; 186 AA.

AC Q9SDR7; 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DE Divergent protein.

OS Forsythia intermedia (Border forsythia).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asterales; euasterids I; Lamiales; Oleaceae; Forsythia.

NCBI_TaxId=55183;
 RN (1)
 RP SEQUENCE FROM N.A.

RC STRAIN=CV; LYWOOD GOLD; TISSUE=STEM;
 RX MEDLINE=99175375; PubMed=10074466;
 RA Gang D.R., Costa M.A., Fujita M., Dinkova-Kostova A.T., Wang H.B.,
 RA Burlat V., Martin W., Sarkkanen S., Davin L.B., Lewis N.G.;

RT "Regiochemical control of monolignol radical coupling: a new paradigm
 for lignin and lignan biosynthesis.";
 RL Chem. Biol. 6:143-151(1999).

DR EMBL: AF210061; AAF25357.1;
 DR InterPro: IPR004265; Disease_resp.
 DR Pfam: PF03018; disease_resp.1.

SQ SEQUENCE 186 AA; 20953 MW; 18F73DBCDA4D17F1 CRC64;
 Query Match 3.7%; Score 7; DB 10; Length 186;
 Best Local Similarity 100.0%; Pred. No. 47;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 GRKPPR 175
 DB 24 GRKPPR 30

Search completed: March 28, 2003, 14:23:22
 Job time : 28.8533 secs

GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:20:52 ; Search time 11.6533 Seconds
(without alignments)
479,722 Million cell updates/sec

Title: US-09-867-274-2

Perfect score: 190
Sequence: 1 OGMQAFKNDATETIPELGEY.....KRPFRASAKANQAELENAY 190

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size : 0

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Issued Patents AA: *
1: /cgn2_6/prodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/iaa/5A_COMB.pep:*
4: /cgn2_6/prodata/1/iaa/5B_COMB.pep:*
5: /cgn2_6/prodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/1/iaa/Backfill1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	190	100.0	213	4	US-09-449-218D-2	Sequence 2, Appli
2	175	92.1	213	4	US-09-449-218D-6	Sequence 6, Appli
3	175	92.1	213	4	US-09-449-218D-8	Sequence 8, Appli
4	128	67.4	213	4	US-09-449-218D-10	Sequence 10, Appli
5	76	40.0	176	4	US-09-449-218D-16	Sequence 16, Appli
6	33	17.4	213	4	US-09-449-218D-14	Sequence 14, Appli
7	32	16.8	211	4	US-09-449-218D-12	Sequence 12, Appli
8	9	4.7	206	1	US-08-468-847B-2	Sequence 2, Appli
9	9	4.7	206	1	US-08-468-847B-20	Sequence 2, Appli
10	9	4.7	206	4	US-09-188-930-159	Sequence 159, App
11	9	4.7	206	4	US-09-188-930-286	Sequence 286, App
12	7	3.7	186	4	US-09-475-316A-13	Sequence 13, Appli
13	7	3.7	917	2	US-08-588-983-16	Sequence 16, Appli
14	7	3.7	917	2	US-08-588-976-16	Sequence 16, Appli
15	6	3.2	32	4	US-09-082-358B-70	Sequence 70, Appli
16	6	3.2	39	3	US-08-554-385-22	Sequence 22, Appli
17	6	3.2	68	2	US-08-836-791-5	Sequence 5, Appli
18	6	3.2	80	1	US-08-011-398B-15	Sequence 15, Appli
19	6	3.2	80	1	US-08-464-051-15	Sequence 15, Appli
20	6	3.2	80	3	US-08-462-498-15	Sequence 15, Appli
21	6	3.2	80	3	US-08-554-385-14	Sequence 14, Appli
22	6	3.2	87	4	US-09-299-495F-31	Sequence 31, Appli
23	6	3.2	88	2	US-08-690-011A-31	Sequence 31, Appli
24	6	3.2	97	2	US-08-245-511-30	Sequence 30, Appli
25	6	3.2	97	2	US-08-600-993A-30	Sequence 30, Appli
26	6	3.2	123	4	US-09-134-001C-3283	Sequence 3283, Ap
27	6	3.2	129	4	US-09-325-932A-50	Sequence 50, Appli

28	6	3.2	182	4	US-09-230-670C-6	Sequence 6, Appli
29	6	3.2	186	1	US-07-960-981-5	Sequence 5, Appli
30	6	3.2	186	5	PCT-US93-09634-5	Sequence 5, Appli
31	6	3.2	187	4	US-09-134-001C-2866	Sequence 2866, Ap
32	6	3.2	199	4	US-09-325-932A-46	Sequence 46, Appli
33	6	3.2	216	2	US-08-821-637-3	Sequence 46, Appli
34	6	3.2	217	4	US-09-134-001C-4620	Sequence 4620, Ap
35	6	3.2	223	4	US-08-928-941D-22	Sequence 22, Appli
36	6	3.2	223	4	US-09-280-590A-22	Sequence 22, Appli
37	6	3.2	224	4	US-09-134-001C-4247	Sequence 4247, Ap
38	6	3.2	229	4	US-09-134-001C-5029	Sequence 5029, Ap
39	6	3.2	243	2	US-08-319-376-2	Sequence 2, Appli
40	6	3.2	248	2	US-08-701-935-3	Sequence 3, Appli
41	6	3.2	248	3	US-09-134-591-3	Sequence 3, Appli
42	6	3.2	254	2	US-08-701-935-6	Sequence 6, Appli
43	6	3.2	254	3	US-09-134-591-6	Sequence 6, Appli
44	6	3.2	255	4	US-09-134-001C-3304	Sequence 3304, Ap
45	6	3.2	269	2	US-08-727-311-3	Sequence 3, Appli

ALIGNMENTS

```
RESULT 1
US-09-449-218D-2
; Sequence 2, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-449-218D-2

Query Match      100.0%; Score 190; DB 4; Length 213;
Best Local Similarity 100.0%; Pred. No. 8.9e-178;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OGMQAFKNDATETIPELGEYPPPELENNKTNRRAENGSRPHHPFTKQVSEYSCREL 60
DB 24 OGMQAFKNDATETIPELGEYPPPELENNKTNRRAENGSRPHHPFTKQVSEYSCREL 83
QY 61 HTRRVVTGPGCSAKPYTELCSGCGGPALLPNAIGRGKMMRPSGPDPRCIPRYRAQR 120
DB 84 HTRRVVTGPGCSAKPYTELCSGCGGPALLPNAIGRGKMMRPSGPDPRCIPRYRAQR 143
QY 121 VOLLCPGSEAPRARKRVRLVASCCKRLTRFNQSELKDFTEARPOGGRKPRPARSAK 180
DB 144 VOLLCPGSEAPRARKRVRLVASCCKRLTRFNQSELKDFTEARPOGGRKPRPARSAK 203
QY 181 ANQAELENAY 190
DB 204 ANQAELENAY 213

RESULT 2
US-09-449-218D-6
; Sequence 6, Application US/09449218D
; Patent No. 6395511
```

```

; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-449-218D-6

Query Match          92.1%; Score 175; DB 4; Length 213;
Best Local Similarity 100.0%; Pred. No. 3.8e-163;
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ELGEPPEPPLENNKTNRAENGRRPPHPPETKDVSEYSCRELHTRVYTGPGCSAK 75
DB 39 ELGEPPEPPLENNKTNRAENGRRPPHPPETKDVSEYSCRELHTRVYTGPGCSAK 98
QY 76 PVTSLVCSGCCPARLLPNAIGRKWRRPSGDPFCIPDRYRAORVOLLCPGGEAPRARK 135
DB 99 PVTSLVCSGCCPARLLPNAIGRKWRRPSGDPFCIPDRYRAORVOLLCPGGEAPRARK 158
QY 136 VRLVASCKCKRLTRFHNSSEIKDFGTDAARPOKGRKRRPPARSAKANOAELENNY 190
DB 159 VRLVASCKCKRLTRFHNSSEIKDFGTDAARPOKGRKRRPPARSAKANOAELENNY 213

RESULT 3
US-09-449-218D-8
; Sequence 8, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-449-218D-8

Query Match          92.1%; Score 175; DB 4; Length 213;
Best Local Similarity 100.0%; Pred. No. 3.8e-163;
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ELGEPPEPPLENNKTNRAENGRRPPHPPETKDVSEYSCRELHTRVYTGPGCSAK 75
DB 39 ELGEPPEPPLENNKTNRAENGRRPPHPPETKDVSEYSCRELHTRVYTGPGCSAK 98
QY 76 PVTSLVCSGCCPARLLPNAIGRKWRRPSGDPFCIPDRYRAORVOLLCPGGEAPRARK 135
DB 99 PVTSLVCSGCCPARLLPNAIGRKWRRPSGDPFCIPDRYRAORVOLLCPGGEAPRARK 135
```

```

DB 99 PVTSLVCSGCCPARLLPNAIGRKWRRPSGDPFCIPDRYRAORVOLLCPGGEAPRARK 158
QY 136 VRLVASCKCKRLTRFHNSSEIKDFGTDAARPOKGRKRRPPARSAKANOAELENNY 190
DB 159 VRLVASCKCKRLTRFHNSSEIKDFGTDAARPOKGRKRRPPARSAKANOAELENNY 213

RESULT 4
US-09-449-218D-10
; Sequence 10, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Cercopithecus pygerythrus
US-09-449-218D-10

Query Match          67.4%; Score 128; DB 4; Length 213;
Best Local Similarity 100.0%; Pred. No. 2.7e-117;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OGMOAFKNDATETIIPELGEYEPPELENNKTNRAENGRRPPHPPETKDVSEYSCREL 60
DB 24 OGMOAFKNDATETIIPELGEYEPPELENNKTNRAENGRRPPHPPETKDVSEYSCREL 83
QY 61 HFTRYVTDGPCRSARPVTELVSQCGCPARLLPNAIGRKWRRPSGDPFCIPDRYRAOR 120
DB 84 HFTRYVTDGPCRSARPVTELVSQCGCPARLLPNAIGRKWRRPSGDPFCIPDRYRAOR 143
QY 121 VOLLCRGG 128
DB 144 VOLLCRGG 151

RESULT 5
US-09-449-218D-16
; Sequence 16, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 176
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-449-218D-16
```

Query Match 40.0%; Score 76; DB 4; Length 176;
Best Local Similarity 100.0%; Pred. No. 1.2e-66;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 53 SEVSCRELHFTRYVTDGPCRSAPVTELVCSGCCGPARLLPNAIGRGKXWRSPGDFRCI 112
DB 45 SEVSCRELHFTRYVTDGPCRSAPVTELVCSGCCGPARLLPNAIGRGKXWRSPGDFRCI 104

QY 113 PDRYRQVRVOLLCPGG 128
DB 105 PDRYRQVRVOLLCPGG 120

RESULT 6
US-09-449-218D-14
Sequence 14, Application US/09449218D
Patent No. 6395511

GENERAL INFORMATION:
APPLICANT: Brunkow, Mary E.
APPLICANT: Galas, David J.
APPLICANT: Kovacevich, Brian
APPLICANT: Mulligan, John T.
APPLICANT: Paepel, Bryan W.
APPLICANT: Van Ness, Jeffrey
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
FILE REFERENCE: 240083.508
CURRENT APPLICATION NUMBER: US/09/449,218D
CURRENT FILING DATE: 1999-11-24
NUMBER OF SEQ ID NOS: 45
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 14
LENGTH: 213
TYPE: PRT
ORGANISM: Rattus norvegicus
US-09-449-218D-14

Query Match 17.4%; Score 33; DB 4; Length 213;
Best Local Similarity 100.0%; Pred. No. 1.3e-24;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 66 VTDGPCRSAPVTELVCSGCCGPARLLPNAIGR 98
DB 89 VTDGPCRSAPVTELVCSGCCGPARLLPNAIGR 121

RESULT 7
US-09-449-218D-12
Sequence 12, Application US/09449218D
Patent No. 6395511

GENERAL INFORMATION:
APPLICANT: Brunkow, Mary E.
APPLICANT: Galas, David J.
APPLICANT: Kovacevich, Brian
APPLICANT: Mulligan, John T.
APPLICANT: Paepel, Bryan W.
APPLICANT: Van Ness, Jeffrey
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
FILE REFERENCE: 240083.508
CURRENT APPLICATION NUMBER: US/09/449,218D
CURRENT FILING DATE: 1999-11-24
NUMBER OF SEQ ID NOS: 45
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 12
LENGTH: 211
TYPE: PRT
ORGANISM: Mus musculus
US-09-449-218D-12

Query Match 16.8%; Score 32; DB 4; Length 211;
Best Local Similarity 100.0%; Pred. No. 1.2e-23;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 TDGPCRSAPVTELVCSGCCGPARLLPNAIGR 98
DB 88 TDGPCRSAPVTELVCSGCCGPARLLPNAIGR 119

RESULT 8
US-08-468-847B-2
Sequence 2, Application US/08468847B
Patent No. 5780263

GENERAL INFORMATION:
APPLICANT: Hastings, Gregg A. and Adams, Mark D.
TITLE OF INVENTION: Human CCN-Like Growth Factor
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
ADDRESSEE: CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,847B
FILING DATE: 6 June 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: MULHINS, J.G.
REGISTRATION NUMBER: 33,073

REFERENCE/DOCKET NUMBER: 325800-442
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 206 AMINO ACIDS
TYPE: AMINO ACID

STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN

US-08-468-847B-2

Query Match 4.7%; Score 9; DB 1; Length 206;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 AFKNDATETI 13
DB 23 AFKNDATETI 31

RESULT 9
US-08-468-847B-20
Sequence 20, Application US/08468847B
Patent No. 5780263

GENERAL INFORMATION:
APPLICANT: Hastings, Gregg A. and Adams, Mark D.
TITLE OF INVENTION: Human CCN-Like Growth Factor
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
ADDRESSEE: CECCHI, STEWART & OLSTEIN

```
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,847B
FILING DATE: 6 June 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: MULLINS, J.G.
REGISTRATION NUMBER: 33,073
REFERENCE/DOCKET NUMBER: 325800-442
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 206 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
US-08-468-847B-20

Query Match          4.7%; Score 9; DB 1; Length 206;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 AFKNDATETI 13
    |||||
Db 23 AFKNDATETI 31

RESULT 10
US-09-188-930-159
; Sequence 159, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 159
; LENGTH: 206
; TYPE: PRT
; ORGANISM: mouse
US-09-188-930-159

Query Match          4.7%; Score 9; DB 4; Length 206;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 AFKNDATETI 13
    |||||
Db 23 AFKNDATETI 31
```

```
RESULT 11
US-09-188-930-286
; Sequence 286, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 286
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Mouse
US-09-188-930-286

Query Match          4.7%; Score 9; DB 4; Length 206;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 AFKNDATETI 13
    |||||
Db 23 AFKNDATETI 31

RESULT 12
US-09-475-316A-13
; Sequence 13, Application US/09475316A
; Patent No. 6210942
; GENERAL INFORMATION:
; APPLICANT: Lewis, No. 6210942man G.
; APPLICANT: Davin, Laurence B.
; APPLICANT: Dinkova-Kostova, Albena T.
; APPLICANT: Fujita, Masayuki
; APPLICANT: Gang, David R.
; APPLICANT: Sarkanen, Simo
; APPLICANT: Ford, Joshua D
; TITLE OF INVENTION: RECOMBINANT PINOSESINOL/LARICRESINOL REDUCTASES,
; TITLE OF INVENTION: RECOMBINANT DIRIGENT PROTEINS AND METHODS OF USE
; FILE REFERENCE: WSUR-1-13793
; CURRENT APPLICATION NUMBER: US/09/475,316A
; CURRENT FILING DATE: 1999-12-30
; PRIOR APPLICATION NUMBER: 09/307,653
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: PCT/US97/20391
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: 60/054,380
; PRIOR FILING DATE: 1997-07-31
; PRIOR APPLICATION NUMBER: 60/030,522
; PRIOR FILING DATE: 1996-11-08
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 186
; TYPE: PRT
; ORGANISM: Forsythia x intermedia
US-09-475-316A-13

Query Match          3.7%; Score 7; DB 4; Length 186;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 GRKRRPR 175
    |||||
Db 24 GRKRRPR 30
```


RESULT 13
US-08-588-983-16
Sequence 16, Application US/08588983
Patent No. 5854067
GENERAL INFORMATION:
APPLICANT: Christopher B. Newgard, et al.
TITLE OF INVENTION: Methods and Compositions
for Inhibiting Hexokinase
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/588,983
FILING DATE: Concurrently herewith
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Fussey, Shelley P.M.
REGISTRATION NUMBER: 39,458
REFERENCE/DOCKET NUMBER: UTSD:424/FUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
TELEX: n/a
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 917 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-588-983-16
Query Match
Best Local Similarity 3.7%; Score 7; DB 2; Length 917;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 47 FETKQVS 53
DB 334 FETKQVS 340
RESULT 14
US-08-588-976-16
Sequence 16, Application US/08588976
Patent No. 5891717
GENERAL INFORMATION:
APPLICANT: Christopher B. Newgard, et al.
TITLE OF INVENTION: Methods and Compositions for
Inhibiting Hexokinase
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/588,976
FILING DATE: Concurrently herewith
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fussey, Shelley P.M.
REGISTRATION NUMBER: 39,458
REFERENCE/DOCKET NUMBER: UTSD:481/FUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
TELEX: n/a
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 917 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-588-976-16

Query Match
Best Local Similarity 3.7%; Score 7; DB 2; Length 917;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 47 FETKQVS 53
DB 334 FETKQVS 340

RESULT 15
US-09-082-358B-70
Sequence 70, Application US/09082358B
Patent No. 6469153
GENERAL INFORMATION:
APPLICANT: Goff, Stephen P.
TITLE OF INVENTION: EIP-1, EIP-3 GENES, ENVELOPE-INTERACTING PROTEINS,
FILE REFERENCE: 0575/54804
CURRENT APPLICATION NUMBER: US/09/082,358B
CURRENT FILING DATE: 1998-05-20
NUMBER OF SEQ ID NOS: 106
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 70
LENGTH: 32
TYPE: PRT
ORGANISM: murine
US-09-082-358B-70

Query Match
Best Local Similarity 3.2%; Score 6; DB 4; Length 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 128 GEAPRA 133
DB 10 GEAPRA 15

Search completed: March 28, 2003, 14:24:27
Job time : 13.6533 secs



GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:22:02 ; Search time 12.16 Seconds
(without alignments)
917.557 Million cell updates/sec

Title: US-09-867-274-2

Perfect score: 190
Sequence: 1 OQWQAFKQDATEIIPELGEY.....KRPFRASAKANQAELENAY 190

Scoring table: OLIGO
Gapex 60.0 , Gapext 60.0

Searched: 237916 seqs, 58723674 residues

Word size : 0

Total number of hits satisfying chosen parameters: 237916

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	190	100.0	190	US-09-867-274-2	Sequence 2, Appli
2	190	100.0	213	US-09-867-274-5	Sequence 5, Appli
3	139	73.2	139	US-09-864-761-47109	Sequence 47109, A
4	32	16.8	185	US-09-867-274-4	Sequence 4, Appli
5	32	16.8	208	US-09-867-274-6	Sequence 6, Appli
6	9	4.7	206	US-09-866-050A-159	Sequence 159, App
7	9	4.7	206	US-09-866-050A-286	Sequence 286, App
8	9	4.7	206	US-09-853-625B-2	Sequence 2, Appli
9	9	4.7	206	US-09-853-625B-10	Sequence 20, Appli
10	8	4.2	183	US-09-867-274-25	Sequence 25, Appli
11	7	3.7	186	US-09-944-160-16	Sequence 16, Appli
12	7	3.7	205	US-09-864-761-36244	Sequence 36244, A
13	7	3.7	373	US-09-738-626-368	Sequence 368, Ap
14	7	3.7	705	US-10-154-386-2	Sequence 2, Appli
15	7	3.7	917	US-09-808-743-10	Sequence 10, Appli
16	6	3.2	46	US-09-925-297-836	Sequence 836, App
17	6	3.2	53	US-10-001-857-164	Sequence 164, App
18	6	3.2	59	US-09-510-332-72	Sequence 72, Appli
19	6	3.2	59	US-09-864-761-34799	Sequence 34799, A

20	6	3.2	87	9	US-10-059-720-31	Sequence 31, Appli
21	6	3.2	87	9	US-10-091-572-334	Sequence 334, App
22	6	3.2	104	10	US-09-864-761-33381	Sequence 33381, A
23	6	3.2	110	10	US-09-741-669-320	Sequence 320, App
24	6	3.2	110	10	US-09-912-020-387	Sequence 387, App
25	6	3.2	110	10	US-09-815-242-10354	Sequence 10354, A
26	6	3.2	110	10	US-09-815-242-13970	Sequence 13970, A
27	6	3.2	112	10	US-09-815-242-5235	Sequence 5235, App
28	6	3.2	113	9	US-10-013-379-17	Sequence 17, Appli
29	6	3.2	117	10	US-09-815-242-12258	Sequence 12258, A
30	6	3.2	117	10	US-09-815-242-12765	Sequence 12765, A
31	6	3.2	124	9	US-09-738-626-5430	Sequence 5430, App
32	6	3.2	124	10	US-09-912-020-250	Sequence 250, App
33	6	3.2	124	10	US-09-815-242-10365	Sequence 10365, A
34	6	3.2	124	10	US-09-815-242-13967	Sequence 13967, A
35	6	3.2	132	9	US-09-981-876-147	Sequence 147, App
36	6	3.2	132	9	US-09-148-545-147	Sequence 147, App
37	6	3.2	137	10	US-09-867-550-832	Sequence 832, App
38	6	3.2	143	10	US-09-815-242-11669	Sequence 11669, A
39	6	3.2	149	10	US-09-764-864-1171	Sequence 1171, App
40	6	3.2	160	10	US-09-815-242-11664	Sequence 11664, A
41	6	3.2	163	9	US-09-925-299-1035	Sequence 1035, App
42	6	3.2	163	10	US-09-925-299-1035	Sequence 1035, App
43	6	3.2	165	9	US-09-738-626-5314	Sequence 5314, App
44	6	3.2	195	10	US-09-858-664A-11	Sequence 11, Appli
45	6	3.2	201	10	US-09-925-300-1646	Sequence 1646, App

ALIGNMENTS

RESULT 1
US-09-867-274-2
Sequence 2, Application US/09867274
Patent No. US20020106650A1
GENERAL INFORMATION: Christopher
APPLICANT: Gao, Yongling
TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
FILE REFERENCE: 01017/37428
CURRENT APPLICATION NUMBER: US/09/867,274
CURRENT FILING DATE: 2001-05-29
PRIOR APPLICATION NUMBER: US 60/208,550
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: US 60/223,542
PRIOR FILING DATE: 2000-08-04
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 190
TYPE: PRT
ORGANISM: Homo sapiens
US-09-867-274-2

Query Match 100.0%; Score 190; DB 10; Length 190;
Best Local Similarity 100.0%; Pred. No. 2.5e-180;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 OQWQAFKQDATEIIPELGEYPPPELENNKTNRAENGSRPHHPFTKQVSESCREL 60
1 OQWQAFKQDATEIIPELGEYPPPELENNKTNRAENGSRPHHPFTKQVSESCREL 60
61 HFTFRVVTGPGCSAKPVELVCSGCCGPARLLPNAIGCKMWRPSGPDPRCIPRYRQOR 120
61 HFTFRVVTGPGCSAKPVELVCSGCCGPARLLPNAIGCKMWRPSGPDPRCIPRYRQOR 120
121 VOLLCPGSAPARRKRVIVASCKCKLRTFRNOSLKDFTGEARPPQKRRPPRASAK 180
121 VOLLCPGSAPARRKRVIVASCKCKLRTFRNOSLKDFTGEARPPQKRRPPRASAK 180
181 ANQAELENAY 190
181 ANQAELENAY 190

```
RESULT 2
US-09-867-274-5
; Sequence 5, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-867-274-5

Query Match      100.0%; Score 190; DB 10; Length 213;
Best Local Similarity 100.0%; Pred. No. 2.7e-180;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGWAFAKNDATETIPELGEYPEPPELENNKTKMRAENGSRPHHPETKDVSEYSCREL 60
DB 24 QGWAFAKNDATETIPELGEYPEPPELENNKTKMRAENGSRPHHPETKDVSEYSCREL 83
QY 61 HFTRYVTDGPCRSKAPVTELVCSGCGPARLLPNAIGRGKWRPSPGDPFRCIPDRYPAQR 120
DB 84 HFTRYVTDGPCRSKAPVTELVCSGCGPARLLPNAIGRGKWRPSPGDPFRCIPDRYPAQR 143
QY 121 VOLLCGGEAPRAKRVLVASCCKRLLTRFHNOSLKDQFGTEARPOGKRRPPRARSAS 160
DB 144 VOLLCGGEAPRAKRVLVASCCKRLLTRFHNOSLKDQFGTEARPOGKRRPPRARSAS 203
QY 181 ANQAELENAY 190
DB 204 ANQAELENAY 213

RESULT 3
US-09-864-761-47109
; Sequence 47109, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
```

```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47109
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC003098.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.62
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.49
; OTHER INFORMATION: SWISSPROT HIT: P45646, EVALU 4.70e-01
US-09-864-761-47109

Query Match      73.2%; Score 139; DB 10; Length 139;
Best Local Similarity 100.0%; Pred. No. 5.1e-130;
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 VSEYSGRELHFTRYVTDGPCRSKAPVTELVCSGCGPARLLPNAIGRGKWRPSPGDPFRC 111
DB 1 VSEYSGRELHFTRYVTDGPCRSKAPVTELVCSGCGPARLLPNAIGRGKWRPSPGDPFRC 60
QY 112 IPDRYPAQRVOLLCGGEAPRAKRVLVASCCKRLLTRFHNOSLKDQFGTEARPOGKRR 171
DB 61 IPDRYPAQRVOLLCGGEAPRAKRVLVASCCKRLLTRFHNOSLKDQFGTEARPOGKRR 120
QY 172 PRPARSAKANQAELENAY 190
DB 121 PRPARSAKANQAELENAY 139

RESULT 4
US-09-867-274-4
; Sequence 4, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 185
```

TYPE: PRT
ORGANISM: Mus musculus
US-09-867-274-4

Query Match 16.8%; Score 32; DB 10; Length 185;
Best Local Similarity 100.0%; Pred. No. 4,4e-24;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 TDGCRSAKPVTELVCSGCGCPARLLPNAIGR 98
DB 65 TDGCRSAKPVTELVCSGCGCPARLLPNAIGR 96

RESULT 5
US-09-867-274-6
Sequence 6, Application US/09867274
Patent No. US20020106650A1
GENERAL INFORMATION:
APPLICANT: Paszty, Christopher
APPLICANT: Gao, Yongming
TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
FILE REFERENCE: 01017/37428
CURRENT APPLICATION NUMBER: US/09/867,274
CURRENT FILING DATE: 2001-05-29
PRIOR APPLICATION NUMBER: US 60/208,550
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: US 60/223,542
PRIOR FILING DATE: 2000-08-04
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn version 3.0
SEQ ID NO 6
LENGTH: 208
TYPE: PRT
ORGANISM: Mus musculus
US-09-867-274-6

Query Match 16.8%; Score 32; DB 10; Length 208;
Best Local Similarity 100.0%; Pred. No. 4.9e-24;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 TDGCRSAKPVTELVCSGCGCPARLLPNAIGR 98
DB 88 TDGCRSAKPVTELVCSGCGCPARLLPNAIGR 119

RESULT 6
US-09-866-050A-159
Sequence 159, Application US/09866050A
Publication No. US20030040471A1
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Onrust, Rene
APPLICANT: Murison, James G.
APPLICANT: Kumble, Krishanand D.
TITLE OF INVENTION: Compositions Isolated From Skin Cells
FILE REFERENCE: 11000.1011c4U
CURRENT APPLICATION NUMBER: US/09/866,050A
CURRENT FILING DATE: 2001-05-24
NUMBER OF SEQ ID NOS: 725
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 159
LENGTH: 206
TYPE: PRT
ORGANISM: Mouse
US-09-866-050A-159

Query Match 4.7%; Score 9; DB 9; Length 206;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 AFKNDATET 13
DB 23 AFKNDATET 31

RESULT 7
US-09-866-050A-286
Sequence 286, Application US/09866050A
Publication No. US20030040471A1
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Onrust, Rene
APPLICANT: Murison, James G.
APPLICANT: Kumble, Krishanand D.
TITLE OF INVENTION: Compositions Isolated From Skin Cells
FILE REFERENCE: 11000.1011c4U
CURRENT APPLICATION NUMBER: US/09/866,050A
CURRENT FILING DATE: 2001-05-24
NUMBER OF SEQ ID NOS: 725
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 286
LENGTH: 206
TYPE: PRT
ORGANISM: Mouse
US-09-866-050A-286

Query Match 4.7%; Score 9; DB 9; Length 206;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 AFKNDATET 13
DB 23 AFKNDATET 31

RESULT 8
US-09-853-625B-2
Sequence 2, Application US/09853625B
Patent No. US20020049304A1
GENERAL INFORMATION:
APPLICANT: Hastings, Gregg A. and Adams, Mark D.
TITLE OF INVENTION: Human CCN-Like Growth Factor
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/853,625B
FILING DATE: 14-May-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/053,587
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: MULLINS, J.G.
REGISTRATION NUMBER: 33,073
REFERENCE/DOCKET NUMBER: 325800-442
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 206 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS: <Unknown>
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-853-625B-2

Query Match 4.7%; Score 9; DB 10; Length 206;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 AFKXDATEI 13
DB 23 AFKXDATEI 31

RESULT 9

US-09-853-625B-20
Sequence 20, Application US/09853625B
Patent No. US20020049304A1
GENERAL INFORMATION:
APPLICANT: Hastings, Gregg A. and Adams, Mark D.
TITLE OF INVENTION: Human CCN-Like Growth Factor
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESSES:
ADDRESSEE: CAREDIA, BYRNE, BAIN, GILFILLIAN,
CECCHI, STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/853,625B
FILING DATE: 14-May-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/053,587
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: MULHINS, J.G.
REGISTRATION NUMBER: 33,073
REFERENCE/DOCKET NUMBER: 325800-442
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 206 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS: <Unknown>
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-853-625B-20

Query Match 4.7%; Score 9; DB 10; Length 206;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 AFKXDATEI 13
DB 23 AFKXDATEI 31

RESULT 10
US-09-867-274-25
Sequence 25, Application US/09867274
Patent No. US20020106650A1
GENERAL INFORMATION:
APPLICANT: Paszty, Christopher
TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
FILE REFERENCE: 01017/37428
CURRENT APPLICATION NUMBER: US/09/867,274
PRIOR FILING DATE: 2001-05-29
PRIOR APPLICATION NUMBER: US 60/208,550
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: US 60/223,542
PRIOR FILING DATE: 2000-08-04
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn version 3.0
SEQ ID NO 25
LENGTH: 183
TYPE: PRT
ORGANISM: Homo sapiens
US-09-867-274-25

Query Match 4.2%; Score 8; DB 10; Length 183;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 FKXDATEI 13
DB 1 FKXDATEI 8

RESULT 11
US-09-944-160-16
Sequence 16, Application US/09944160
Patent No. US20020174452A1
GENERAL INFORMATION:
APPLICANT: Lewis, No. US20020174452A1man
APPLICANT: Davin, Laurence
TITLE OF INVENTION: Monocot Seeds with Increased Lignan
FILE REFERENCE: WSRU117983
CURRENT APPLICATION NUMBER: US/09/944,160
CURRENT FILING DATE: 2001-08-30
PRIOR APPLICATION NUMBER: US 60/230,632
PRIOR FILING DATE: 2000-09-07
NUMBER OF SEQ ID NOS: 49
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16
LENGTH: 186
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: dirigent amino acid sequence from plasmid pAP1244
US-09-944-160-16

Query Match 3.7%; Score 7; DB 9; Length 186;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 169 GRKPRPR 175
DB 24 GRKPRPR 30

RESULT 12
US-09-864-761-36244
Sequence 36244, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.
 APPLICANT: Chen, Wensheng
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
 FILE REFERENCE: Aemica-X-1
 CURRENT APPLICATION NUMBER: US/09/864,761
 CURRENT FILING DATE: 2001-05-23
 PRIOR APPLICATION NUMBER: US 60/180,312
 PRIOR FILING DATE: 2000-02-04
 PRIOR APPLICATION NUMBER: US 60/207,456
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: US 09/632,366
 PRIOR FILING DATE: 2000-08-03
 PRIOR APPLICATION NUMBER: GB 24263,6
 PRIOR FILING DATE: 2000-10-04
 PRIOR APPLICATION NUMBER: US 60/236,359
 PRIOR FILING DATE: 2000-09-27
 PRIOR APPLICATION NUMBER: PCT/US01/00666
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00667
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00664
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00669
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00665
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00668
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00663
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00662
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00661
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00670
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: US 60/234,687
 PRIOR FILING DATE: 2000-09-21
 PRIOR APPLICATION NUMBER: US 09/608,408
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: US 09/774,203
 PRIOR FILING DATE: 2001-01-29
 NUMBER OF SEQ ID NOS: 49117
 SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 SEQ ID NO 36244
 LENGTH: 205
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: MAP TO AL050305.9
 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.93
 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.99
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.99
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.92
 OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.9
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
 OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.3
 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.96
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.9
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
 OTHER INFORMATION: EST HUMAN HIT: A181773.1, EVALUATE 1.00e-17
 OTHER INFORMATION: SWISSPROT HIT: Q09472, EVALUATE 2.40e-01
 US-09-864-761-36244

Query Match 3.7%; Score 7; DB 10; Length 205;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 11 TRIPPEL 17
 Db 145 TRIPPEL 151

RESULT 13
 US-09-738-626-3568
 Sequence 3568, Application US/09738626
 Publication No. US20020197605A1
 GENERAL INFORMATION:
 APPLICANT: NAKAGAWA, SATOSHI
 APPLICANT: MIZOGUCHI, HIROSHI
 APPLICANT: ANDO, SEIKO
 APPLICANT: HAYASHI, MIKIRO
 APPLICANT: OCHIAI, KEIKO
 APPLICANT: YOKOI, HARUHIKO
 APPLICANT: TATEISHI, NAOKO
 APPLICANT: SENOH, AKIHIRO
 APPLICANT: IKEDA, MASATO
 APPLICANT: OZAKI, AKIO
 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 FILE REFERENCE: 249-125
 CURRENT APPLICATION NUMBER: US/09/738,626
 CURRENT FILING DATE: 2000-12-18
 PRIOR APPLICATION NUMBER: JP 99/377484
 PRIOR FILING DATE: 1999-12-16
 PRIOR APPLICATION NUMBER: JP 00/159162
 PRIOR FILING DATE: 2000-04-07
 PRIOR APPLICATION NUMBER: JP 00/280988
 PRIOR FILING DATE: 2000-08-03
 NUMBER OF SEQ ID NOS: 7059
 SOFTWARE: PatentIn ver. 3.0
 SEQ ID NO 3568
 LENGTH: 373
 TYPE: PRT
 ORGANISM: Corynebacterium glutamicum
 US-09-738-626-3568

Query Match 3.7%; Score 7; DB 9; Length 373;
 Best Local Similarity 100.0%; Pred. No. 44;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 63 TRYVTDG 69
 Db 39 TRYVTDG 45

RESULT 14
 US-10-154-386-2
 Sequence 2, Application US/10154386
 Publication No. US20030026793A1
 GENERAL INFORMATION:
 APPLICANT: Angiogene Inc.
 APPLICANT: Guy, Louis-Georges
 TITLE OF INVENTION: HIPPOXIA INDUCING FACTORS AND USES THEREOF FOR INDUCING ANGIOGENESIS
 TITLE OF INVENTION: IMPROVING MUSCULAR FUNCTIONS
 FILE REFERENCE: 5600-81
 CURRENT APPLICATION NUMBER: US/10/154,386
 CURRENT FILING DATE: 2002-05-23
 PRIOR APPLICATION NUMBER: US 60/292,630
 PRIOR FILING DATE: 2001-05-22
 PRIOR APPLICATION NUMBER: US 60/354529
 PRIOR FILING DATE: 2002-02-08
 NUMBER OF SEQ ID NOS: 13
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 2
 LENGTH: 705
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-154-386-2

Query Match 3.7%; Score 7; DB 9; Length 705;
 Best Local Similarity 100.0%; Pred. No. 77;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 172 PRPRARS 178

Db 556 PRPRARS 562

RESULT 15

US-09-808-743-10
; Sequence 10, Application US/09808743
; Patent No. US20020068711A1
; GENERAL INFORMATION:
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE
; APPLICANT: PEDERSEN, Peter
; APPLICANT: MATHUPALA, Saroj
; TITLE OF INVENTION: ARREST OF PROLIFERATION OF HIGHLY GLYCOLYTIC TUMORS
; FILE REFERENCE: JHU1720-1
; CURRENT APPLICATION NUMBER: US/09/808,743
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: US 60/189,222
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 917
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-808-743-10

Query Match 3.7%; Score 7; DB 10; Length 917;

Best local Similarity 100.0%; Pred. No. 97;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 FETKQVS 53
|||||

Db 334 FETKQVS 340

Search completed: March 28, 2003, 14:24:57
Job time : 14.16 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 28, 2003, 17:27:08 ; Search time 1369.11 Seconds
(without alignments)
7523.378 Million cell updates/sec

Title: US-09-867-274-3

Perfect score: 636
Sequence: 1 atgcagccctcactagcccc.....agctggagaaagcctcactag 636

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_estbm:*
3: em_estbn:*
4: em_estbu:*
5: em_estbv:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_ges:*
18: em_ges_hum:*
19: em_ges_inv:*
20: em_ges_pin:*
21: em_ges_vrt:*
22: em_ges_fun:*
23: em_ges_mam:*
24: em_ges_mus:*
25: em_ges_other:*
26: em_ges_pro:*
27: em_ges_tod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	636	100.0	1990	11 AK017295	AK017295 Mus muscu
2	608	95.6	627	10 BB636457	BB636457 Mus muscu
3	583	91.7	663	10 BB637315	BB637315 BB637315
4	469.4	73.8	511	10 BB638050	BB638050 BB638050
5	391.4	61.5	419	10 BB101082	BB101082 UI-R-B11-
6	336.6	52.9	360	9 A1556282	A1556282 UI-R-C2p-

C	7	289.4	45.5	315	9 A1113131	A1113131 UI-R-C2p-
C	8	273.4	43.0	291	10 BB111224	BB111224 UI-R-B11-
C	9	138.8	21.8	535	17 A0171546	A0171546 HS 3088 B
C	10	136.2	21.4	254	12 BB523030	BB523030 UI-R-C2p-
	11	83.2	13.1	1612	11 AK002396	AK002396 Mus muscu
	12	83.2	13.1	1613	11 AK007893	AK007893 Mus muscu
	13	83.2	13.1	1690	11 AK002240	AK002240 Mus muscu
	14	83.2	13.1	1691	11 AK007967	AK007967 Mus muscu
	15	82.2	12.9	669	12 BB607657	BB607657 MYL 00059
	16	80.4	12.6	472	12 BR420024	BR420024 UI-R-B22-
	17	80.4	12.6	532	12 BF407514	BF407514 UI-R-B22-
	18	79.4	12.5	385	10 AM434705	AM434705 UI-R-B2p
	19	78.2	12.3	789	11 AK007935	AK007935 Mus muscu
	20	76.4	12.0	687	13 B1081999	B1081999 602879824
	21	76	11.9	623	13 B1143761	B1143761 602907224
	22	74.4	11.7	577	13 B1682960	B1682960 464177 MA
	23	74.2	11.7	584	12 BF041721	BF041721 BP25005B
	24	73.4	11.7	925	17 CNS0091P	AL053013 Drosoph11
	25	73.4	11.5	355	9 AA197904	AA197904 mv05c05.r
	26	73.4	11.5	433	9 AA067619	AA067619 fa19b09.s
	27	73	11.5	320	9 AA065771	AA065771 Drosoph11
	28	71.4	11.2	925	17 CNS0091P	AL053013 Drosoph11
C	29	70.6	11.1	627	17 A2876400	A2876400 2M0191D21
	30	68.4	10.8	683	12 BG705527	BG705527 602685416
	31	68.2	10.7	1020	13 B1767517	B1767517 603061406
	32	68	10.7	932	17 CNS00720	AL066742 Drosoph11
	33	67.2	10.6	715	13 B1489720	B1489720 603032217
	34	67.2	10.6	735	13 B1763021	B1763021 603047812
	35	67.2	10.6	740	13 B1760787	B1760787 603043808
	36	67.2	10.6	850	9 A1524992	A1524992 A1524992
	37	67.2	10.6	1028	14 B0067667	B0067667 AGENCOURT
	38	67.2	10.6	1094	14 BM924116	BM924116 AGENCOURT
	39	67.2	10.6	1111	13 BM548167	BM548167 AGENCOURT
	40	67.2	10.6	1873	11 AF361494	AF361494 Homo sapi
	41	66.4	10.4	597	14 BM751691	BM751691 K-EST0027
	42	65.2	10.3	1076	14 BM921437	BM921437 AGENCOURT
	43	64.8	10.2	578	14 BM709058	BM709058 UI-E-CQ1-
	44	64.8	10.2	618	14 BM748765	BM748765 K-EST0023
	45	64.8	10.2	645	14 BM835363	BM835363 K-EST0110

ALIGNMENTS

RESULT 1	AK017295	1990 bp	mRNA	linear	HTC 19-JAN-2002
LOCUS	AK017295				
DEFINITION	Mus musculus 6 days neonate head cDNA, RIKEN full-length enriched library, clone:5430411E23:sclerostin, full insert sequence.				
ACCESSION	AK017295				
VERSION	AK017295.1 GI:12856464				
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (Strain: C57BL/6J) 6 days neonate head cDNA to mRNA, clone_11b:RIKEN full-length enriched mouse cDNA library				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.				
AUTHORS	Carninci, P. and Hayashizaki, Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE	2				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PUBMED	11042159				
REFERENCE	3				

OY	541	GGGCGAGACCGCGCCGCCGCAGAAGGTGCAAGCCCGCGC3CCCGGGAGACC	600	
Db	584	GGCGCGAGACCGCGCCGCCGCCAAGAGGTGCAAGCCCGCGC3CCCGGGAGACC	643	
OY	601	AAAGCCAACGAGCGGAGCTGAGAGAACGCTTACTAG	636	
Db	644	AAAGCCAACGAGCGGAGCTGAGAGAACGCTTACTAG	679	
RESULT 2				
LOCUS	B636457			
DEFINITION	B636457 RIKEN full-length enriched, adult male aorta and vein Mus musculus cDNA A530001M15 5', mRNA sequence.	667 bp	linear EST 26-Oct-2001	
ACCESSION	B636457			
VERSION	B636457.1	GI:16472326		
KEYWORDS	EST.			
SOURCE	Mus musculus			
ORGANISM	house mouse.			
REFERENCE	Emmaliya; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 667)			
AUTHORS	Arakawa,T., Carrinci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,U., Komoto,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.			
TITLE	RIKEN Mouse ESTs (Arakawa,T., et al. 2001)			
JOURNAL	Unpublished (2001)			
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel.: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-resesc.riken.go.jp, url:http://genome.gsc.riken.go.jp/ Carrinci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Komoto,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) wagii,K., Fujiwake,S., Inoue,K., Togawa,Y., Itawa,M., Ohara,E., Watabiki,M., Yoneda,Y., Ichikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,U., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Komoto,H., Fukunishi,Y., Shibata,K., Itoh,M., Carrinci,P., Sugahara,Y. and Hayashizaki,Y. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) Kondo,S., Shinagawa,A., Saito,T., Kiyoawa,H., Yamanake,I., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawaji,J., Shibata,K. and Hayashizaki,Y. Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001) Please visit our web site (http://genome.gsc.riken.go.jp) for further details. e mouse tissues.			
FEATURES				
SOURCE	Location/Qualifiers			
	1..667			
	/organism="Mus musculus"			
	/db_xref="taxon:10090"			
	/clone="A530001M15"			
	/clone_id="RIKEN full-length enriched, adult male aorta and vein"			
	/sex="male"			

	/tissue_type="aorta and vein"
	/dev_stage="adult"
	/lab_host="DH10B"
	/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'
	GAGGAGAAGATCCACAGACTCTTTTNTTTTTTTTTTNVN 3']
	cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGGAGAAGATTTCGATTAAATAAATTCccccccccc 3']
	cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLX I."
BASE COUNT	119 A 234 C 212 G 100 T 2 others
ORIGIN	
Query Match	95.6%; Score 608; DB 10; Length 667;
Best Local Similarity	98.1% ; Pred. No. 4.le-121;
Matches 614; Conservative	0; Mismatches 12; Indels 0; Gaps 0;
OY	1 ATGCAGGCCCTCACTAGAGCCCCGGTGCTCATCTCGCCTATTGTGACGGCTGCTGTGCT 60
Dd	41 ATGCAGGCCCTCACTAGAGCCCCGGTGCTCATCTCGCCTATTGTGACGGCTGCTGTGCT 100
OY	61 GTGAGGGCCAGGGGTGCGAACCTTCAGAAATGAATGCACAGAGTCATCCAGGGCTT 120
Dd	101 GTGAGGGCCAGGGGTGCGAACCTTCAGAAATGAATGCACAAAGTCATCCAGGGCTT 160
OY	121 GAGAGTACCCCAGACCCTCTCTGAGAACAAACAGACAATGAACCGGGCGAGAAATGA 180
Dd	161 GAGAGTACCCCAGACCCTCTCTGAGAACAAACAGACAATGAACCGGGCGAGAAATGA 220
OY	181 GGACAGACCTCCCAACCATCTTAGACGCGCAAAGATGTGTCAGTAAGATGCGCGAG 240
Dd	221 GGACAGACCTCCCAACCATCTTAGACGCGCAAAGATGTGTAAGTAAGATGCGCGAG 280
OY	241 CTGCATACACCCGTTTCTTGAAGAGCGGCCCATATGCCGAGGCCAGCCGATCACCGAG 300
Dd	281 CTGCATACACCCGTTTCTTGAAGAGCGGCCCATATGCCGAGGCCAGCCGATCACCGAG 340
OY	301 TTGGTGCTCCGGCAGATGCGGCCCGCGGGCGGTGCTGCGCCAACGCATCGGGGCGGTG 360
Dd	341 TTGGTGCTCCGGCAGATGCGGCCCGCGGGCGGTGCTGCTGCCAACGCATCGGGGCGGTG 400
OY	361 AAGTGTTGGCCCGGAACGGAACCGGATTTCGGTGCATCCCGGATCGCTACCGCGCGAG 420
Dd	401 AAGTGTTGGCCCGGAACGGAACCGGATTTCGGTGCATCCCGGATCGCTACCGCGCGAG 460
OY	421 CGGGTGCAGCTGTGTGCCCCGGGGGGCGCGCGCGCGCTTGGCGAAAGTGCCTTGGTG 480
Dd	461 CGGGTGCAGCTGTGTGCCCCGGGGGGCGCGNCGCGCGCTTGGCGAAAGTGCCTTGGTG 520
OY	481 GCCTGTCAAGTCAAGCAAGCGCTCACCCGCTTTCACAACAACAGTGCAGACTCAAGACTTC 540
Dd	521 GCCTGTCAAGTCAAGCAAGCGCTCACCCGCTTTCACAACAACAGTGCAGACTTCAGAGACTTC 580
OY	541 GGGCGGAGACCGCGCGCGCGCGCAAGAGGTGCAAGCCGCGCGCCCGCGCGCGGGGAGCC 600
Dd	581 GGGCGGAGAAAACCGCGCGCGCGCGCAAAATGTGCAAGCCGCGCGCGCGCGGGGAGACC 640
OY	601 AAAGCAACACAGCGCGAGCTGGAGAA 626
Dd	641 AAAGCAACACAGCGCGAGCTGGAGAA 666

	BEB637315	
LOCUS	BEB637315	623 bp mRNA linear EST 26-OCT-2001
DEFINITION	BEB637315 RIKEN full-length enriched, adult male aorta and vein Mus musculus cDNA clone A530065C22.5', mRNA sequence.	
ACCESSION	BEB637315	
KEYWORDS	BEB637315.1 GI:16473100	
SOURCE	EST.	
ORGANISM	Mus musculus house mouse. Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
REFERENCE	1 (bases 1 to 623)	
AUTHORS	Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,U., Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okasaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.	
TITLE	RIKEN Mouse ESTs (Arakawa,T., et al. 2001)	
JOURNAL	Unpublished (2001)	
COMMENT	Contact: Yoshinide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Shuto-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel.: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gscc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/ Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) Wegli,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanabe,K., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,U., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Konno,H., Fukuishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamataka,I., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and Hayashizaki,Y. Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001) Please visit our web site (http://genome.gsc.riken.go.jp) for further details. e mouse tissues.	
FEATURES	Location/Qualifiers	
Source	1.. 623 /organism="Mus musculus" /db_xref="taxon:10090" /clone="A530065C22" /clone_jib="RIKEN full-length enriched, adult male aorta and vein" /sex="male" /tissue_type="aorta and vein" /dev_stage="adult" /lab_host="DH10B" /note="Site_1: Saili; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'	

BASE COUNT	106 a	222 c	199 g	96 t
ORIGIN				
Query Match	91.7%;	Score 583;	DB 10;	Length 623;
Best Local Similarity	100.0%;	Pred. No. 9.7e-116;		
Matches 583; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	GAGGACAGAAGATCCAAAGCCTTTTGTTCCTTTTNN 3'), cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATTCTCGATTATTAATTAATCCCCCCC 3'] cDNA was cleaved with XhoI and BamH. Vector: a modified pBluescript KS(+) after bulk excision from Lambda			
Db	FLC I."			
QY	1	ATGCAGCCCTCACTAGCCCCGGTGCTCATCTGCCACTTTGTGACGCTGCTTCTGTGCT	60	
Db	41	ATGCAGCCCTCACTAGCCCCGGTGCTCATCTGCCACTTTGTGACGCTGCTTCTGTGCT	100	
QY	61	GTGAGGGCCAGAGGGGTGSCAAGCTTTCAGGAATGATGCACAGAGTATCCAGGGCTT	120	
Db	101	GTGAGGGCCAGAGGGGTGSCAAGCTTTCAGGAATGATGCACAGAGTATCCAGGGCTT	160	
QY	121	GGAAGTACCCTCCAGACCTCTCTTAGAACAACGACCATGAAACCGGGGAGAGAAATGA	180	
Db	161	GGAAGTACCCTCCAGACCTCTCTTAGAACAACGACCATGAAACCGGGGAGAGAAATGA	220	
QY	181	GGCACACCTCCCAACATCCCTATGACGCCCAAAGATGTGTGCGAGTACAGCTGCCCGAG	240	
Db	221	GGCACACCTCCCAACATCCCTATGACGCCCAAAGATGTGTGCGAGTACAGCTGCCCGAG	280	
QY	241	CTGCACATAACCCGCTTCTTGA CAGACGGCCCATCCGACGCGCACCGGCTACCGAG	300	
Db	281	CTGCACATAACCCGCTTCTTGA CAGACGGCCCATCCGACGCGCACCGGCTACCGAG	340	
QY	301	TTCGTGTGCTCCGGGCAAGTGGGGCCCCCGGGGCTGTGCTGCCATCGGGGGCGTG	360	
Db	341	TTCGTGTGCTCCGGGCAAGTGGGGCCCCCGGGGCTGTGCTGCCATCGGGGGCGTG	400	
QY	361	AAGTGTGGCGCCGCAACGGAACCGGATTTCCGCTCATCCCGATCCGATCCGCGCGCAG	420	
Db	401	AAGTGTGGCGCCGCAACGGAACCGGATTTCCGCTCATCCCGATCCGATCCGCGCGCAG	460	
QY	421	CGGGTGCAGCTGCTGTGCCCGGGGGCGCGCGCGCTCGCGCAAGTGTGCTGTGGT	480	
Db	461	CGGGTGCAGCTGCTGTGCCCGGGGGCGCGCGCGCTCGCGCAAGTGTGCTGTGGT	520	
QY	481	GCTTGTGTCAAGTGCNAAGCGCTCAACCCGCTTCCAACACAGTCCGACTTCAGGAATTC	540	
Db	521	GCTTGTGTCAAGTGCNAAGCGCTCAACCCGCTTCCAACACAGTCCGACTTCAGGAATTC	580	
QY	541	GGGCGCGAGAACCGCGCGCGCGCAGAAAGGTGCGCAAGCGCGCGC	583	
Db	581	GGGCGCGAGAACCGCGCGCGCGCAGAAAGGTGCGCAAGCGCGCGC	623	
RESULT 4				
BB638050				
LOCUS	BB638050	511 bp	mRNA	linear EST 31-AUG-2001
DEFINITION	BB638050 RIKEN full-length enriched, 3 days neonate thymus Mus musculus cDNA A630006E24 5', mRNA sequence.			
ACCESSION	BB638050			
VERSION	BB638050.1 GI:15401086			
KEYWORDS	EST.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 511)			
AUTHORS	Arkawa,T., Carinci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,U., Komno,H., Koude,			


```

/db_xref="taxon:10116"
/clone="UI-R-BJ1-acy-h-10-0-UI"
/clone_1b="UI-R-BJ1"
/lab_host="DH10B (Life Technologies)"
/Note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-BJ1
library is a subtracted library derived from the following
tissues: atrium at 16.5 dpc, ventricle at 16.5 dpc, AV
canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc,
AV canal at 15 dpc, ventricle at 13 dpc, and adult heart.
For a detailed description of the library from which this
clone was derived, please visit our web site at
ratseq.eng.uiowa.edu. The subtraction has been previously
described in (Bonaldo, Lennon and Soares, Genome Research
6:791-806, 1996)
TAG_SEQ=None found"
BASE COUNT      56 a      135 c      155 g      73 t
ORIGIN

```

```

Query Match      61.5%; Score 391.4; DB 10; Length 419;
Best Local Similarity 96.2%; Pred. No. 1.7e-74;
Matches 401; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 145 GAGAACACAGACATGAACCGGGCGAGATGAGGACAGACCTCCACACATCCCTAT 204
DB 417 GAGAACACAGACATGAACCGGGCGAGATGAGGACAGACCTCCACACATCCCTAT 358
QY 205 GAGCCCAAGATGTGTCCGATACAGCTGCCGCGAGCTGCATACACCGCTTCTGACA 264
DB 357 GACACCAAGATGTGTCCGATACAGCTGCCGCGAGCTGCATACACCGCTTCTGACC 298
QY 265 GAGGAGCCATCCGAGGAGCGGACCGGATCAACGATTTGGATGCTCCGAGCGAGC 324
DB 297 GAGGAGCCATCCGAGGAGCGGACCGGATCAACGATTTGGATGCTCCGAGCGAGC 238
QY 325 CCCGCGGCTGCTGCCCAACCGCCATCGGCGCGTGAAGTGTGCGCCCGGAACGACG 384
DB 237 CCCGCGGCTGCTGCCCAACCGCCATCGGCGCGTGAAGTGTGCGCCCGGAACGACG 178
QY 385 GATTTCGCTGCATCCGAGATCGCTACCGCGCGAGCGGGTGCAGCTGTGTGCCCCG 444
DB 177 GACTTCGCTGCATCCGAGATCGCTACCGCGCGAGCGGGTGCAGCTGTGTGCCCCG 118
QY 445 GGGCGGGCGCGGCTGCGCGCAAGTGCGTGTGTGCTCGTGCAGATGCAAGCGCTC 504
DB 117 GGGCGGGCGCGGCTGCGCGCAAGTGCGTGTGTGCTCGTGCAGATGCAAGCGCTC 58
QY 505 ACCGCTTCCACCAACAGTCCGAGCTCAAGACTTTCGGGCGGAGACCGCGCGCG 561
DB 57 ACCGCTTCCACCAACAGTCCGAGCTCAAGACTTTCGGGCGGAGACCGCGCGCG 1

```

```

RESULT 6
A1556282/c      360 bp      mRNA      linear      EST 23-MAR-1999
LOCUS
DEFINITION
UI-R-C2p-rk-a-03-0-UI s1 UI-R-C2p Ratnus norvegicus cDNA clone
ACCESSION
A1556282
VERSION
A1556282.1 GI:4488645
KEYWORDS
EST.
SOURCE
Norway rat.
Rattus norvegicus
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 360)
AUTHORS
Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE
Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL
Genome Res. 6 (9), 791-806 (1996)
MEDLINE
9704477
COMMENT
Contact: Soares, MB
Program for Rat Gene Discovery and Mapping

```

```

FEATURES
source
    University of Iowa
    451 Eckstein Medical Research Building Iowa City, IA 52242, USA
    Tel: 319 335 8250
    Fax: 319 335 9555
    Email: msoares@blue.weeg.uiowa.edu
    Oligo-dT track not found, Not I site shown in beginning of sequence
    is likely internal to the message. cDNA library preparation: M.
    Fatima Bonaldo, Ph.D. Clone distribution: clones will be available
    through Research Genetics (www.resgen.com) The following repetitive
    elements were found in this cDNA sequence: 97-125,
    >GC-rich#low-complexity
    Seq primer: M13 Forward.
    Location/Qualifiers
        1..360
        /organism="Rattus norvegicus"
        /strain="Sprague-Dawley"
        /db_xref="taxon:10116"
        /clone="UI-R-C2p-rk-a-03-0-UI"
        /clone_1b="UI-R-C2p"
        /dev_stage="adult"
        /lab_host="DH10B (Life Technologies)"
        /note="Vector: pT73D-Pac (Pharmacia) with a modified
        polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-C2p
        library is a subtracted library derived from the UI-R-C1
        library, which is a subtracted library derived from the
        UI-R-C0 library. The UI-R-C0 library consisted of a
        mixture of individually tagged normalized libraries
        constructed from rat placenta, adult lung, brain, liver,
        kidney, heart, spleen, ovary, muscle, 8, 12 and 18-day
        embryo. The tag is a string of 3-5 nucleotides present
        between the Not I site and the oligo-dT track which allows
        identification of the library of origin of a clone within
        the mixture. The subtracted library (UI-R-C2p) was
        constructed as follows: PCR amplified cDNA inserts from
        UI-R-C1 clones from which 3' ESTs had been derived was
        used as a driver in a hybridization with the UI-R-C1
        library in the form of single-stranded circles. The
        remaining single-stranded circles (subtracted library) was
        purified by hydroxyapatite column chromatography,
        converted to double-stranded circles and electroporated
        into DH10B bacteria (Life Technologies) to generate the
        UI-R-C2p library. This procedure has been previously
        described (Bonaldo, Lennon and Soares, Genome Research 6:
        791-806, 1996)"

```

```

BASE COUNT      54 a      120 c      131 g      54 t      1 others
ORIGIN
Query Match      52.9%; Score 336.6; DB 9; Length 360;
Best Local Similarity 95.8%; Pred. No. 1e-62;
Matches 345; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 202 TATGACGCCAAGATGTGTCCGATGACGCTGCCGAGCTGCATACCCGCTTCTG 261
DB 360 TATGACGCCAAGATGTGTCCGATGACGCTGCCGAGCTGCATACCCGCTTCTG 301
QY 262 ACAGACGCGCATGCGAGAGCGCAAGCGGATCGATGATGCTCGGCGAGTGC 321
DB 300 ACAGACGCGCATGCGAGAGCGCAAGCGGATCGATGATGCTCGGCGAGTGC 241
QY 322 GGGCGCGCGGCTGCTGCCCAACGCCATCGGCGCGCTGAAGTGTGCGCCGCAACG 381
DB 240 GGGCGCGCGGCTGCTGCCCAACGCCATCGGCGCGCTGAAGTGTGCGCCGCAACG 181
QY 382 CGGATTTCCGCTGCATCCCGATGCTACCGCGCGAGCGGCTGCTGTGCCCC 441
DB 180 CGGATTTCCGCTGCATCCCGATGCTACCGCGCGAGCGGCTGCTGTGCCCC 121
QY 442 GGGGGGCGGCGCGGCTCGCGCAAGTGTGCTGAGGCTGCTGCTGAATGCAAGCG 501
DB 120 GGGGGGCGGCGCGGCTCGCGCAAGTGTGCTGAGGCTGCTGCTGAATGCAAGCG 61
QY 502 CTCACCGCTTCCACCAACAGTCCGAGCTCAAGACTTTCGGGCGGAGACCGCGCGCG 561

```


CC and designated as Cloaked-2, derived from human and mouse. The cloaked-2
 CC polypeptides can be expressed by standard recombinant methodology. The
 CC cloaked-2 polynucleotides are useful in gene therapy and antisense
 CC therapy. The cloaked-2 polypeptides and polynucleotides are useful for
 CC treating, preventing, ameliorating or detecting diseases and disorders of
 CC the kidney (e.g. anemia, hypertension or low blood pressure), heart (e.g.
 CC cardiac hypertrophy, congestive heart failure, myocardial infarction,
 CC arrhythmias, atherosclerosis, hypertension or low blood pressure),
 CC skeletal muscle (e.g. muscular dystrophy or cachexia), placenta (e.g.
 CC congenital abnormalities or miscarriage), liver (e.g. hepatitis or
 CC cirrhosis), pancreas (e.g. diabetes or pancreatitis), thyroid (e.g.
 CC Grave's disease or myxedema) or adrenal cortex (e.g. Cushing's disease
 CC or Addison's disease), homeostasis or metabolic diseases (e.g. obesity,
 CC cancer or myopathies), infections, or autoimmune diseases. Selective
 CC binding agents may be used to modulate the biological activities of
 CC cloaked-2 polypeptides or to detect cloaked-2 polypeptide levels in a
 CC sample. Transgenic non-human animals are useful for drug candidate
 CC screening. The present sequence represents the mouse cloaked-2 mature
 CC polypeptide.

CC Sequence 188 AA;

Query Match 90.7%; Score 951; DB 23; Length 188;
 Best Local Similarity 89.5%; Pred. No. 4.6e-84;
 Matches 170; Conservative 10; Mismatches 8; Indels 2; Gaps 1;

QY 1 OGMQAFKNDATETIIPELGEYEPPELENNKTNRAENGSRPHHPFTKVSEYSCEL 60
 Db 1 OGMQAFKNDATETIIPGLGEYEPPEPP--ENNQTNRRAENGSRPHHPYDAKDVSEYSCEL 58
 QY 61 HFTRYVTGPGCRSAKPVTELVCSGCCGPALLPNAIGRKWMPSPGPFRCIPDRYRQR 120
 Db 59 HFTRYVTGPGCRSAKPVTELVCSGCCGPALLPNAIGRKWMPSPGPFRCIPDRYRQR 118
 QY 121 VQLLCGGEAPRARKVRLVASCCKRLTRFNOSLKDFTGEARPOKGRKPRRARSK 180
 Db 119 VQLLCGGEAPRARKVRLVASCCKRLTRFNOSLKDFTGEARPOKGRKPRRARSK 178
 QY 181 ANQAELENAY 190
 Db 179 ANQAELENAY 188

RESULT 14
 AAY96432
 ID AAY96432 standard; Protein; 211 AA.

AC AAY96432;

DT 12-SEP-2000 (first entry)

DE Murine TGF-beta binding protein (BBER).

KM osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KM BBER; gene therapy; antisense therapy; fracture; bone mineralization.

OS Mus musculus.

PN WO200032773-A1.

PD 08-JUN-2000.

PF 24-NOV-1999; 99WO-US27990.

PR 27-NOV-1998; 98US-0110283.

PA (DARN-) DARWIN DISCOVERY LTD.

PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepker BW;
 PI Van Ness J, Winkler DG;

XX WPI, 2000-412321/35.
 XX N-PSDB; AAA29058.
 DR

XX Nucleic acids (1) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures
 XX
 PS Claim 5; Page 124; 162pp; English.

CC This shows a murine transforming growth factor-beta (TGF-beta)
 CC binding protein designated mBER. The cDNA and protein may be used for
 CC prevention, treatment and diagnosis of diseases associated with
 CC inappropriate BBER expression. For example, they may be used to treat
 CC disorders associated with decreased TGF-beta BP expression. The cDNA or
 CC vectors may be administered to treat diseases by rectifying mutations or
 CC expressions in a patient's genome that affect the activity of BBER by
 CC expressing inactive proteins or to supplement the patients own production
 CC of BBER polypeptides. The nucleic acids may be used for recombinant
 CC production of BBER, gene therapy, antisense therapy, as probes for
 CC diagnostic assays and for functional studies. BBER may be used to raise
 CC antibodies and for identification of BBER modulators. BBER antagonists
 CC may be used to increase bone mineral content for the treatment of
 CC disorders such as osteopenia, osteoporosis, fractures and other disorders
 CC associated with low mineral content.

CC Sequence 211 AA;

Query Match 90.7%; Score 951; DB 21; Length 211;
 Best Local Similarity 89.5%; Pred. No. 5.2e-84;
 Matches 170; Conservative 10; Mismatches 8; Indels 2; Gaps 1;

QY 1 OGMQAFKNDATETIIPELGEYEPPELENNKTNRAENGSRPHHPFTKVSEYSCEL 60
 Db 24 OGMQAFKNDATETIIPGLGEYEPPEPP--ENNQTNRRAENGSRPHHPYDAKDVSEYSCEL 81
 QY 61 HFTRYVTGPGCRSAKPVTELVCSGCCGPALLPNAIGRKWMPSPGPFRCIPDRYRQR 120
 Db 82 HFTRYVTGPGCRSAKPVTELVCSGCCGPALLPNAIGRKWMPSPGPFRCIPDRYRQR 141
 QY 121 VQLLCGGEAPRARKVRLVASCCKRLTRFNOSLKDFTGEARPOKGRKPRRARSK 180
 Db 142 VQLLCGGEAPRARKVRLVASCCKRLTRFNOSLKDFTGEARPOKGRKPRRARSK 201
 QY 181 ANQAELENAY 190
 Db 202 ANQAELENAY 211

RESULT 15
 ABB07210
 ID ABB07210 standard; Protein; 211 AA.

AC ABB07210;

DT 26-MAR-2002 (first entry)

DE Mouse cloaked-2 polypeptide sequence.

KM Cloaked-2; cysteine knot motif; nephrotropic; cardiac; immunomodulator;
 KM hepatotropic; antiinflammatory; antihypertoid; cytoprotective;
 KM antianemic; hypotensive; antiarrhythmic; antiarteriosclerotic; muscular;
 KM antidiabetic; anorectic; gene therapy; cell therapy; antisense therapy;
 KM mouse.

OS Mus musculus.

XX

XX

XX

XX

XX

FT Key Location/Qualifiers
 FT Peptide 1..23
 FT Protein /note= "signal peptide"
 FT Protein 24
 FT Misc-difference /note= "mature protein (ABB07208)"
 FT /note= "this region is missing in the sequence provided
 FT in the sequence listing but has been indicated
 FT correctly in the sequence in the Figure"

```
XX MO200192308-A2.
PN
XX 06-DEC-2001.
PD
XX 29-MAY-2001; 2001WO-US17478.
PF
XX 01-JUN-2000; 2000US-208550P.
PR 04-AUG-2000; 2000US-223542P.
XX
PA (AMGE-) AMGEN INC.
XX
PI Paszty CJ, Gao Y;
XX
XX WPI; 2002-114325/15.
DR N-PSDB; ABA94294.
XX
XX New human and mouse cysteine-knot polypeptide designated as Cloaked-2,
PT for treating or preventing kidney, heart (e.g. myocardial infarction)
PT or liver (e.g. hepatitis) diseases
XX
XX Example 2; Fig 2; 170pp; English.
PS
XX The invention relates to polypeptides comprising a cysteine knot motif
CC and designated as Cloaked-2, derived from human and mouse. The cloaked-2
CC polypeptides can be expressed by standard recombinant methodology. The
CC cloaked-2 polynucleotides are useful in gene therapy and antisense
CC therapy. The cloaked-2 polypeptides and polynucleotides are useful for
CC treating, preventing, ameliorating or detecting diseases and disorders of
CC the kidney (e.g. anemia, hypertension or low blood pressure), heart (e.g.
CC cardiac hypertrophy, congestive heart failure, myocardial infarction,
CC arrhythmias, atherosclerosis, hypertension or low blood pressure),
CC skeletal muscle (e.g. muscular dystrophy or cachexia), placenta (e.g.
CC congenital abnormalities or miscarriage), liver (e.g. hepatitis or
CC cirrhosis), pancreas (e.g. diabetes or pancreatitis), thyroid (e.g.
CC Grave's disease or myxedema) or adrenal cortex (e.g. Cushing's disease
CC or Addison's disease), homeostasis or metabolic diseases (e.g. obesity,
CC cancer or myopathies), infections, or autoimmune diseases. Selective
CC binding agents may be used to modulate the biological activities of
CC cloaked-2 polypeptides or to detect cloaked-2 polypeptide levels in a
CC sample. Transgenic non-human animals are useful for drug candidate
CC screening. The present sequence represents the mouse cloaked-2
CC polypeptide.
XX
SQ Sequence 211 AA;
Query Match 90.7%; Score 951; DB 23; Length 211;
Best Local Similarity 89.5%; Pred. No. 5.2e-84;
Matches 170; Conservative 10; Mismatches 8; Indels 2; Gaps 1;
QY 1 QGWAFAKNDATETIPELGEYEPPELLENNKTKRAENGGRPHHPETKDVSEYSCREL 60
Db 24 QGWAFAKNDATETIPELGEYEPPELLENNKTKRAENGGRPHHPETKDVSEYSCREL 81
QY 61 HFTRYVTDGPGCRSAKPYTELVCSCGCGPARLLPNAIGRKWMPSPDPFRICIPRYRAQR 120
Db 82 HFTRYVTDGPGCRSAKPYTELVCSCGCGPARLLPNAIGRKWMPSPDPFRICIPRYRAQR 141
QY 121 VQLICPGGEARPRARKRLVASCCKRLTRFNQSELKDPGTEARPOKGRKPRPRASAK 180
Db 142 VQLICPGGEARPRARKRLVASCCKRLTRFNQSELKDPGTEARPOKGRKPRPRASAK 201
QY 181 ANOAELENAY 190
Db 202 ANOAELENAY 211
```

Search completed: March 28, 2003, 14:16:13
Job time : 54.2 secs

GenCore version 5.1.4 p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:13:46 ; Search time 27.36 Seconds
(without alignments)
667.600 Million cell updates/sec

Title: US-09-867-274-2
Perfect score: 1049
Sequence: 1 QGMQAFKNDATIEIIPELGEV.....KPRPRASAKANQALENAV 190

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 10%
Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	270	25.7	134	2 T08710	hypothetical prote
2	94.5	9.0	159	2 T51373	luteinizing hormon
3	87.5	8.3	158	2 A61091	luteotropin beta chai
4	87.5	8.3	601	2 T22025	hypothetical prote
5	87.5	8.3	601	2 D89711	protein F40E10.4 (
6	82.5	7.9	1469	2 B3665	slit protein 2 pre
7	81	7.7	270	2 S71793	head-inducing fact
8	80.5	7.7	715	2 T12534	hypothetical prote
9	80	7.6	1480	2 A3665	slit protein 1 pre
10	79	7.5	118	2 PN0141	luteotropin beta chai
11	79	7.5	178	2 T51824	tumor-suppressive
12	79	7.5	178	2 A47291	probable finger pr
13	79	7.5	866	2 T29197	hypothetical prote
14	78.5	7.5	166	2 T51242	luteinizing hormon
15	78.5	7.5	969	2 A70912	probable leus prot
16	78	7.4	1142	2 T30272	hypothetical prote
17	78	7.4	1210	1 GQHUE	epidermal growth f
18	77.5	7.4	1042	2 A57534	mucin 5AC (clone L
19	77	7.3	118	2 PN0139	luteotropin beta chai
20	77	7.3	510	2 A42750	insulinoma-associat
21	77	7.3	520	2 S47142	matting type A prot
22	77	7.3	839	2 T50590	class I INCEMP pro
23	76.5	7.3	1223	2 S29717	adenylate cyclase
24	76	7.2	398	2 A35281	intergenetary muc
25	76	7.2	2142	2 B35098	MHC class III hist
26	75.5	7.2	403	2 C70832	hypothetical prote
27	75.5	7.2	646	2 T02398	hypothetical prote
28	75.5	7.2	1436	2 A46496	antigen Wc1.1 prec
29	75	7.1	2153	2 T30074	hypothetical prote

30	74.5	7.1	206	2 G87592	hypothetical prote
31	74.5	7.1	824	2 T10020	leucine-cRNA ligas
32	74.5	7.1	972	2 T10023	leucine-cRNA ligas
33	74.5	7.1	1063	2 T03743	bifocal protein -
34	74	7.1	269	2 E69381	hypothetical prote
35	74	7.1	343	2 G35070	apolipoprotein H-r
36	74	7.1	702	2 E72775	probable helicase
37	74	7.1	877	2 T50591	class II INCEMP pr
38	74	7.1	2090	2 T30075	hypothetical prote
39	73.5	7.0	682	2 A44493	serum-inducible kl
40	73.5	7.0	864	2 H85335	hypothetical prote
41	73.5	7.0	864	2 T04518	hypothetical prote
42	73	7.0	155	2 C95002	hypothetical prote
43	73	7.0	301	2 WMBEP9	UL49 protein - hum
44	73	7.0	613	2 T00077	96g-like protein -
45	73	7.0	1589	2 T42233	submaxillary mucin

ALIGNMENTS

RESULT 1

T08710
hypothetical protein DKFP564D206.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 13-Aug-1999
C:Accession: T08710
R:Mandut, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
Submitted to the Protein Sequence Database, March 1999
A:Reference number: Z16471
A:Accession: T08710
A:Molecule type: mRNA
A:Residues: 1-134 <N&M>
A:Cross-references: EMBL:AL050024
A:Experimental source: fetal brain; clone DKFP564D206
C:Genetics:
A>Note: DKFP564D206.1

Query Match 25.7% Score 270; DB 2; Length 134;

Best local similarity 44.2%; Pred. NO. 4e-17; Mismatches 57; Conservative 23; Mismatches 43; Indels 6; Gaps 4;

QY	57	CRELHFRVYVDGPRSAKPYTELVCSGQCGPARLPNALIGRG--KWM-RPSGDPDFRCI 112	
DB	3	CRELRSTYISDQCTISPLKELVCAEGELLPLPWIGGVTGKXMSRSSQEMRCV 62	
QY	113	PDRYRAQVOLLCPGEAPRARKVRLVASCKKRLTRFHNOSELKDGTEAARPOKGRKP 172	
DB	63	NDKRTTORIQLOCDG-STRTYKITVTVACKCKRYTROHNSHNFSPSPAKPVQHRE 121	
QY	173	RPRA-RSAK 180	
DB	122	KRRASKSK 130	

RESULT 2

T51373
luteinizing hormone beta subunit - turkey
C:Species: Meleagris gallopavo (common turkey)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
C:Accession: T51373
R:Yon, S.; Foster, L.K.; Slibby, J.L.; el Halawani, M.E.; Foster, D.N.
J. Mol. Endocrinol. 14, 117-129, 1995
A>Title: Sequence analysis of the turkey LH beta subunit and its regulation by gonadotro
A:Reference number: T51373; MUID:95290073; PMID:7772235
A:Accession: T51373
A:Molecule type: translated from GB/EMBL/DBJ
A:Residues: 1-159 <YOU>
A:Cross-references: GB:L35519; NID:g530952; PIDN:AAA74125.1; PID:g530953
C:Genetics:
A:Gene: LH-beta
C:Superfamily: pituitary glycoprotein hormone beta chain

Query Match 9.0%; Score 94.5; DB 2; Length 159;
Best Local Similarity 29.4%; Pred. No. 0.25; Indels 21; Gaps 5;
Matches 32; Conservative 7; Mismatches 49;

Qy 39 GGRPHHPEFKDVSEVSCRELFHFRVYTDGPCRSAPVTELVSGGCGPARLLPNAIGR 98
Db 43 GGRPP-----CRPINTVAVEKDECPQCMATTTTACGGYCRFR-----EPPVR 85

Qy 99 GKWMRPSGPDPRCPIDRYRAQVOLL-CPGGEAPRARKVRLVASCCKR 146
Db 86 SPLGRP--POSSCTYGAALRYERMAWMCPTIGSDPRV-LLPVALSRCRCAR 131

RESULT 3
A61091
Nucleotide beta chain precursor - chicken
N/Alternate names: luteinizing hormone beta chain
C/Species: Gallus gallus (chicken)
C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 08-Dec-1995
R/Accession: A61091
R/Note: T. J. Ando, H. Ueda, T. Kubokawa, K.; Higashinakagawa, T.; Ishii, S.
J. Mol. Endocrinol. 3, 129-137, 1989
A/Title: Molecular cloning and nucleotide sequence analysis of the putative cDNA for the
A/Reference number: A61091; MUID:89374710; PMID:2476156
A/Accession: A61091
A/Status: not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-158 <NOC>
C/Superfamily: pituitary glycoprotein hormone beta chain
C/Keywords: glycoprotein; hormone; pituitary
F/1.39/Domains: signal sequence #status predicted <SIG>
F/48-73,62-96,65-127,77-149,111-139,129-132/Disulfide bonds: #status predicted
F/52/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 8.3%; Score 87.5; DB 2; Length 158;
Best Local Similarity 27.9%; Pred. No. 1;
Matches 31; Conservative 8; Mismatches 47; Indels 25; Gaps 5;

Qy 39 GGRPHHPEFKDVSEVSCRELFHFRVYTDGPCRSAPVTELVSGGCGPARLLPNAIGR 98
Db 43 GGRPP-----CRPINTVAVEKDECPQCMATTTTACGGYCRFR-----EPPVR 85

Qy 99 GKWMRPSG--PDRFCIPDRYRAQVOLL-CPGGEAPRARKVRLVASCCKR 146
Db 84 --YKSLGPPPSACTYGAALRYERMAWMCPTIGSDPRV-LLPVALSRCRCAR 131

RESULT 4
T22025
Hypothetical protein F40E10.4 - Caenorhabditis elegans (fragment)
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C/Accession: T22025
R/Smyle, R.
Submitted to the EMBL Data Library, February 1996
A/Reference number: Z19503
A/Accession: T22025
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-601 <WLL>
A/Cross-references: EMBL:Z69792; PIDN:CAA93668.1; GSPDB:GN00028; CESP:F40E10.4
C/Experimental source: clone F40E10
C/Genetics:
A/Gene: CESP:F40E10.4
A/Map position: X

Query Match 8.3%; Score 87.5; DB 2; Length 601;
Best Local Similarity 20.6%; Pred. No. 3.9;
Matches 21; Conservative 17; Mismatches 45; Indels 19; Gaps 2;

Qy 44 HHPEFKDVSEVSCRELFHFRVYTDGPCRSAPVTELVSGGCGPARLLPNAIGRKWMR 103
Db 54 --GEONCCTAIVKVKRKVMTCCKNG-TTKISTVHIIRQCQE 592

RESULT 5
D89711
Protein F40E10.4 [imported] - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Nov-2001
C/Accession: D89711
R/Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A/Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A/Reference number: A75000; MUID:99069613; PMID:9851916
A/Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C_eleg
A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A/Accession: D89711
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-601 <STO>
A/Cross-references: GB:chr_X; PIDN:CAA93668.1; PID:93877014; GSPDB:GN00028; CESP:F40E10.4
C/Genetics:
A/Gene: F40E10.4
A/Map position: X

Query Match 8.3%; Score 87.5; DB 2; Length 601;
Best Local Similarity 20.6%; Pred. No. 3.9;
Matches 21; Conservative 17; Mismatches 45; Indels 19; Gaps 2;

Qy 44 HHPEFKDVSEVSCRELFHFRVYTDGPCRSAPVTELVSGGCGPARLLPNAIGRKWMR 103
Db 54 --GEONCCTAIVKVKRKVMTCCKNG-TTKISTVHIIRQCQE 592

RESULT 6
B3665
Silt protein 2 precursor - fruit fly (Drosophila melanogaster)
C/Species: Drosophila melanogaster
C/Date: 30-Apr-1991 #sequence_revision 30-Apr-1991 #text_change 02-Aug-2002
C/Accession: B3665
R/Rothberg, J.M.; Jacobs, J.R.; Goodman, C.S.; Artavanis-Tsakonas, S.
Genes Dev. 4, 2169-2187, 1990
A/Title: silt: an extracellular protein necessary for development of midline glia and con
A/Reference number: B3665; MUID:91099665; PMID:2176636
A/Accession: B3665
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-1469 <ROT>
A/Cross-references: GB:X53959
C/Genetics:
A/Gene: FlyBase:silt
A/Cross-references: FlyBase:FBgn0003425
C/Superfamily: fruit fly silt protein; EGF homology; leucine-rich alpha-2-glycoprotein re
F/66-91/Domains: proteoglycan amino-terminal homology <PAH1>
F/101-124/Domains: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F/125-148/Domains: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F/149-172/Domains: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F/173-196/Domains: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F/197-220/Domains: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F/228-313/Domains: proteoglycan carboxyl-terminal homology <PCSI1>
F/323-346/Domains: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
F/347-370/Domains: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
F/371-394/Domains: leucine-rich alpha-2-glycoprotein repeat homology <LRR8>
F/395-418/Domains: leucine-rich alpha-2-glycoprotein repeat homology <LRR9>
F/419-442/Domains: leucine-rich alpha-2-glycoprotein repeat homology <LRR10>

Query Match 8.3%; Score 87.5; DB 2; Length 601;
Best Local Similarity 20.6%; Pred. No. 3.9;
Matches 21; Conservative 17; Mismatches 45; Indels 19; Gaps 2;

Qy 44 HHPEFKDVSEVSCRELFHFRVYTDGPCRSAPVTELVSGGCGPARLLPNAIGRKWMR 103
Db 54 --GEONCCTAIVKVKRKVMTCCKNG-TTKISTVHIIRQCQE 592

Query Match 8.3%; Score 87.5; DB 2; Length 601;
Best Local Similarity 20.6%; Pred. No. 3.9;
Matches 21; Conservative 17; Mismatches 45; Indels 19; Gaps 2;

F/450-494/Domain: proteoglycan carboxyl-terminal homology <PCS2>
 F/512-537/Domain: proteoglycan amino-terminal homology <PAH3>
 F/547-571/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR11>
 F/572-595/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR13>
 F/596-619/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR13>
 F/620-643/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR14>
 F/651-695/Domain: proteoglycan carboxyl-terminal homology <PCS3>
 F/708-733/Domain: proteoglycan amino-terminal homology <PAH4>
 F/743-766/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR15>
 F/767-790/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR16>
 F/846-890/Domain: proteoglycan carboxyl-terminal homology <PCS4>
 F/1028-1061/Domain: EGF homology <EGF2>
 F/1068-1099/Domain: EGF homology <EGF2>
 F/1115-1148/Domain: EGF homology <EGF1>

Query Match 7.9%; Score 82.5; DB 2; Length 1469;
 Best Local Similarity 20.0%; Pred. No. 27;
 Matches 32; Conservative 23; Mismatches 62; Indels 43; Gaps 6;

QY 4 QAFKNDATETIIPELGEYPEPP-PELENNKTMRAENGRR-PPHAPETK-----50
 DB 1333 QSEBEDQDFMDETPHKEPVPDCL-NCRRSGRCVPSNARJGCKCKHGR 1387
 QY 51 -----DVESEYCRELHFRVYTDGCRSAKPVTELVCSGGCPARLLPNAIGRGKMRPSC 106
 DB 1388 GRYCDQASTCRKEQVREYTYENDCRSROPKAKCVGGG-----1428
 QY 107 PDRFCIPDRYRAQVOLLCPGSEAPRAKRVLVASCCKR 146
 DB 1429 -NCCCAKIVRRKRVVVC-SNNKRYIKNDIYRKGCGCK 1466

RESULT 7
 S71793
 head-inducing factor Cerberus - African clawed frog
 C/Species: Xenopus laevis (African clawed frog)
 C/Date: 12-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 21-Jul-2000
 C/Accession: S71793
 R/Bowmeester, T.; Kim, S.H.; Sasaki, Y.; Lu, B.; de Robertis, E.M.
 Nature 382, 595-601, 1996
 A/Title: Cerberus is a head-inducing secreted factor expressed in the anterior endoderm
 A/Reference number: S71793; MUID:96318220; PMID:8757128
 A/Accession: S71793
 A/Status: preliminary; nucleic acid sequence not shown
 A/Status: preliminary; nucleic acid sequence not shown
 A/Molecule type: mRNA
 A/Residues: 1-270 <BOU>
 A/Cross-references: EMBL:U64831; NID:G1513087; PIND:AAC6012.1; PID:G1513088

Query Match 7.7%; Score 81; DB 2; Length 270;
 Best Local Similarity 17.7%; Pred. No. 6.8;
 Matches 28; Conservative 34; Mismatches 70; Indels 26; Gaps 6;
 QY 4 QAFKNDATETIIPELGEYPEPPPELENNKTMRAENGRRPHHPEFTKDVSEVSCRELHPT 63
 DB 122 RSPDKXNTET-----EKGCACFMNNFLVKGANGAPQNTSHSKAOELMKCKTLPT 175
 QY 64 RYVTDGCRSAKPVTELVCSGGCPARLLPNAIGRGKMRPSCDPR-----CIPRYRA 118
 DB 176 QNIVHENC-DRMYIQNNLCFGKISLHV-----PNOQDRRNTCSHCLPSKFTL 222
 QY 119 QRVOLLCPGSEAPRAKRVLVASCCK-RLTRFHNOSE 155
 DB 223 NHLLNCTGSKNV-VKVMVMEBCTEAKHNSPHQTAQ 259

RESULT 8
 T12534
 hypothetical protein DKFP434B094.1 - human (fragment)
 C/Species: Homo sapiens (man)
 C/Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 02-Sep-2000
 C/Accession: T12534
 R/Mamuti, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, June 1999

A/Reference number: 217524
 A/Accession: T12534
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-715 <MAN>
 A/Cross-references: EMBL:AL080149
 A/Experimental source: adult testis; clone DKFP434B094
 C/Genetics:
 A/Note: DKFP434B094.1
 C/Superfamily: bromodomain homology
 F/113-168/Domain: bromodomain homology <BRO>

Query Match 7.7%; Score 80.5; DB 2; Length 715;
 Best Local Similarity 23.9%; Pred. No. 20;
 Matches 49; Conservative 20; Mismatches 87; Indels 49; Gaps 9;

QY 2 GQAFKNDATETIIPELGEY-PEPPPELENNKTM--NRAENGRRPHHPEFTKDVSEVSC 57
 DB 294 GLEGFEEDGALGPENGEGDCKSPKLEPSDALPLPSNSETNSEPP-----TLKPELNPE 349
 QY 58 RELHFRVYTDGCRSAKPVTELVCSGGCPARLLPNAIGRGKMRPSCDPRCIPDRYR 117
 DB 350 GSKLFRKRVTFDSEHSACTQSALV-SG-----RPEPTTASSGDVPA 390
 QY 118 AQRVOLLCPGSEAPRAKRVLVASCCKRLTRFHNOSELQGTGAARPOKGR-----171
 DB 391 AAASAVAEPAQSVNRTSVLF--CKSKSVS--PPKSAKTETQPTSPQGTGTFLSV 444
 QY 172 -----PPRARSARANOAELE 187
 DB 445 LPRLETLQPRRSRST-CGDSEVE 468

RESULT 9
 A36665
 slit protein 1 precursor - fruit fly (Drosophila melanogaster)
 C/Species: Drosophila melanogaster
 C/Date: 30-Apr-1991 #sequence_revision 30-Apr-1991 #text_change 02-Aug-2002
 C/Accession: A36665; A31640; S13523
 R/Rothberg, J.M.; Jacobs, J.R.; Goodman, C.S.; Artavanis-Tsakonas, S.
 Genes Dev. 4, 2169-2187, 1990
 A/Title: slit: an extracellular protein necessary for development of midline glia and con
 A/Reference number: A36665; MUID:91099665; PMID:2176636
 A/Accession: A36665
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-1480 <ROT>
 A/Cross-references: GB:X53959; NID:98614; PIND:CAA37910.1; PID:98615
 R/Rothberg, J.M.; Hartley, D.A.; Walther, Z.; Artavanis-Tsakonas, S.
 Cell 55, 1047-1059, 1988
 A/Title: slit: An EGF-homologous locus of D. melanogaster involved in the development of
 A/Reference number: A31640; MUID:89077533; PMID:3144436
 A/Accession: A31640
 A/Molecule type: DNA
 A/Residues: 881-1182, 'G', '1185-1404', 'GT', '1463-1464', 'YHA', 'RO2>
 A/Cross-references: GB:M23543; NID:G340939; PID:G514357
 C/Genetics:
 A/Gene: FlyBase:slit
 A/Cross-references: FlyBase:FBgn0003425
 A/Introns: 1351/3
 C/Superfamily: fruit fly slit protein; EGF homology; leucine-rich alpha-2-glycoprotein re
 C/Keywords: alternative splicing; growth factor
 F/66-91/Domain: proteoglycan amino-terminal homology <PAH1>
 F/101-124/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR1>
 F/125-148/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR2>
 F/149-172/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR3>
 F/173-196/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR4>
 F/197-220/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR5>
 F/228-272/Domain: proteoglycan carboxyl-terminal homology <PCS1>
 F/288-313/Domain: proteoglycan amino-terminal homology <PAH2>
 F/323-346/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR6>
 F/347-370/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR7>
 F/371-394/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR8>

F/395-418/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR9>
 F/419-442/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR10>
 F/450-494/Domain: proteoglycan carboxyl-terminal homology <PC8>
 F/512-537/Domain: proteoglycan amino-terminal homology <PAH>
 F/547-571/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR11>
 F/572-595/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR12>
 F/596-619/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR13>
 F/620-643/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR14>
 F/651-695/Domain: proteoglycan carboxyl-terminal homology <PC8>
 F/708-733/Domain: proteoglycan amino-terminal homology <PAH>
 F/743-766/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR15>
 F/767-790/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR16>
 F/791-814/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR17>
 F/815-838/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LR18>
 F/846-890/Domain: proteoglycan carboxyl-terminal homology <PC8>
 F/1028-1061/Domain: EGF homology <EGF>
 F/1068-1099/Domain: EGF homology <EGF2>
 F/1115-1148/Domain: EGF homology <EGF1>

Query Match 7.5%; Score 80; DB 2; Length 1480;
 Best Local Similarity 18.0%; Pred. No. 45;
 Matches 31; Conservative 23; Mismatches 62; Indels 56; Gaps 5;

QY 4 QAEKNDATETIPELGEYRPP--PELEN-----NKTWN 34
 DB 1333 QEEBDDQDMDEPHIKERPVDCLEKCRGRSCVPSNNDGYCKKHGRGYCD 1392
 QY 35 RAENGCPHPHPETKDVSEYSCRELFHTYVDGPRSAKPYTELVSQCCPALLRN 94
 DB 1393 QGREGSTPP-----TVTAISTCKKEQVREYTYENDCRSHQPKYAKCVGCG----- 1439
 QY 95 AIGRGKMWRSRGPDRFCIPDRYRAORVOLLCPGGEAPRAKRVLVASCCKR 146
 DB 1440 -----NCCCAKIVRRKRVAMVC-SNNRKYITNLDIVKCGCTK 1477

RESULT 10

PN0141
 Luteolin beta chain - sperm whale
 N/Alternate names: luteinizing hormone beta chain
 C/Species: Physter catodon (sperm whale)
 C/Date: 07-May-1993 #sequence_revision 07-May-1993 #text_change 07-May-1999
 C/Accession: PN0141
 R/Pankov, Y.A.; Karasev, V.S.
 Biochimia 49, 1004-1018, 1984
 A/Title: Luteinizing hormone of the sperm-whale: amino acid sequence of reduced and carb
 A/Reference number: PN0141; PMID:84281133; PMID:6466737
 A/Accession: PN0141
 A/Molecule type: protein
 A/Residues: 1-118 <PAN>
 A/Note: article in Russian with English abstract
 C/Superfamily: pituitary glycoprotein hormone beta chain
 C/Keywords: glycoprotein; hormone
 F/9-34,23-57,26-88,38-110,72-100,90-93/Disulfide bonds: #status predicted
 F/13/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 7.5%; Score 79; DB 2; Length 118;

Best Local Similarity 27.6%; Pred. No. 4.5;
 Matches 34; Conservative 14; Mismatches 55; Indels 20; Gaps 7;

QY 57 CRELHFRYVTDGPRSAKPYTELVSQCCGPARLLPNAIGRGKMWRSRGPDRFCIPDR 115
 DB 9 CRFINATLANONZACPCITFTTISICAGCPSMVRVLPAL-----PVPSPVCTYRQ 61
 QY 116 YRAORVOLLCPGGEAPRAKRVLVASCCK--RLTFPHNOSLKFDTAARPK-GRK 171
 DB 62 LRFASIRLPGCPGVNBMV-SFVALSCHGCPRLSS-----SDCGPRAQPLACNRS 113

QY 172 PRP 174
 DB 114 PRP 116

RESULT 11

151824
 tumor-suppressive gene - rat
 C/Species: Rattus sp. (rat)
 C/Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 05-Nov-1999
 C/Accession: 151824
 R/Sakiyama, S.; Ozaki, T.; Enomoto, H.
 Adv. Enzyme Regul. 34, 247-255, 1994
 A/Title: Molecular cloning and characterization of a cDNA showing tumor-suppressive acti
 A/Reference number: 151824; PMID:95028756; PMID:7942277
 A/Accession: 151824
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-178 <RES>
 A/Cross-references: GB:572637; NID:9619290; PIDN:AAB32215.1; PID:9619291

Query Match 7.5%; Score 79; DB 2; Length 178;
 Best Local Similarity 25.0%; Pred. No. 6.8;
 Matches 38; Conservative 18; Mismatches 68; Indels 28; Gaps 9;

QY 57 CRELHFRYVTDGPRSAKPYTELVSQCCGPARLLPNAIGRGKMWRSRGPDR-----RC 111
 DB 34 CEAKITQIVHSGC-EKSKIQNPAICUGQCF-SYSPVNTF-----POSTESLVHCDSC 84
 QY 112 IPDRYRAORVOLLCPG-GEAPRAK-RVLVASCCKRLTFPHNOSLKF--GTEARPK 167
 DB 85 MPAQSMWEIVTLCEGCHHEVPRVDKLVKIVHSCQACGKEPSHEGLNVYMGEDGGSQ 144
 QY 168 KG---KKRPPRAKSAK-----ANQAELENA 189
 DB 145 PSHSHSHPHGQCPTPEPEPPGAQVEEEGA 176

RESULT 12

A47291
 Probable finger protein - rat
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
 C/Accession: A47291
 R/Ozaki, T.; Sakiyama, S.
 Proc. Natl. Acad. Sci. U.S.A. 90, 2593-2597, 1993
 A/Title: Molecular cloning and characterization of a cDNA showing negative regulation in
 A/Reference number: A47291; PMID:93219329; PMID:8385338
 A/Accession: A47291
 A/Status: preliminary
 A/Molecule type: nucleic acid
 A/Residues: 1-178 <OZA>
 A/Cross-references: GB:X66872; NID:9296927; PIDN:CAA47344.1; PID:9296928
 A/Experimental source: fibroblast 3Y1 cells
 A/Note: sequence extracted from NCBI backbone (NCBIN:128602, NCBI:P:128603)
 C/Keywords: DNA binding; zinc finger

Query Match 7.5%; Score 79; DB 2; Length 178;

Best Local Similarity 25.0%; Pred. No. 6.8;
 Matches 38; Conservative 18; Mismatches 68; Indels 28; Gaps 9;

QY 57 CRELHFRYVTDGPRSAKPYTELVSQCCGPARLLPNAIGRGKMWRSRGPDR-----RC 111
 DB 34 CEAKITQIVHSGC-EKSKIQNPAICUGQCF-SYSPVNTF-----POSTESLVHCDSC 84
 QY 112 IPDRYRAORVOLLCPG-GEAPRAK-RVLVASCCKRLTFPHNOSLKF--GTEARPK 167
 DB 85 MPAQSMWEIVTLCEGCHHEVPRVDKLVKIVHSCQACGKEPSHEGLNVYMGEDGGSQ 144

QY 168 KG---KKRPPRAKSAK-----ANQAELENA 189
 DB 145 PSHSHSHPHGQCPTPEPEPPGAQVEEEGA 176

RESULT 13

T29197
 hypothetical protein T03F1.9 - Caenorhabditis elegans
 C/Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C/Accession: T29197

R/du, Z.: Le, T.T.

Submitted to the EMBL Data Library, February 1997
A/Description: The sequence of C. elegans cosmid T03F1.

A/Reference number: Z20586

A/Accession: T29197

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-866 <DUZ>

A/Cross-references: EMBL:U88169, PIDN:AA84237.1, GSPDB:GN00019, CESP:T03F1.9

A/Experimental source: strain Bristol N2; clone T03F1

A/Genetics:

A/Map position: 1

A/Intons: 15/3; 84/3; 153/3; 194/3; 523/2; 710/1, 760/1

Query Match 7.5%; Score 79; DB 2; Length 866;

Best Local Similarity 21.7%; Pred. No. 33;

Matches 53; Conservative 36; Mismatches 85; Indels 70; Gaps 12;

QY 7 KNDATETIPLGEY-----PEPPELENNKTMRAENG---GRPHHPETKDYSEYSC 57

DB 621 EDATETETKRSNGGVYDGEADSNRSINIBLALTVNEBPAHDSALDPEDDT 680

QY 58 RELHFRVYTDGFCRSAPK-----VTGLVSGGCGPARLLPNAIGKMWKPS 105

DB 681 NEM---RSSDEEMERAPPRKSSRRRLGLSDSLGLASSRRRPNDTFDETWTYBE 737

QY 106 GPDR--CIPDRYRAQV---OLLCPEGAP---PARKVRL--VAS----- 141

DB 738 -PNSKGNRRRTTRKMKLKEHQLMKP--EDAPDGVRRSTRVRVPSWLGEGVYVNSPI 795

QY 142 CKKRLT-----RFNQSGLKDPFGTEARPPQKRRPPRAASACA 181

DB 796 SGCKRLGTAVVVKDPRLCYRTADVTRATERELDKAKRALAQEKQQRONASGR 855

QY 182 NOAE 185

DB 856 HDS 859

RESULT 14

151242

luteinizing hormone beta-subunit - quail

C/Species: Coturnix coturnix (quail)

C/Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999

C/Accession: 151242

R/Ando, H.; Ishii, S.

Gen. Comp. Endocrinol. 93, 357-368, 1994

A/Title: Molecular cloning of complementary deoxyribonucleic acids for the pituitary gly-

urnix coturnix japonica).

A/Reference number: 151241; MUID:94252550; PMID:7515015

A/Accession: 151242

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-166 <AND>

A/Cross-references: GB:S70834, NID:9546921, PIDN:AA830867.1, PID:9546922

C/Superfamily: pituitary glycoprotein hormone beta chain

Query Match 7.5%; Score 78.5; DB 2; Length 166;

Best Local Similarity 26.4%; Pred. No. 7;

Matches 29; Conservative 9; Mismatches 47; Indels 25; Gaps 5;

QY 40 GRPHHPETKDYSEYSCRELHFTRYTDPGCRSAKPVTELVSGGCGPARLLPNAIGRG 99

DB 52 GRPP-----CRPINVTVAVEKECPQCAVTTACGTCRTREPV----- 91

QY 100 KMWKPSG--PDPRCIPDRYRAQVOLL--CPGGEAPRARKRLVLAACKCR 146

DB 92 -YRSPGPPQSSCTYGALRYERMDLWGCPIGSDPKV--ILPVALSRCAR 139

RESULT 15

A70912

probable leus protein - Mycobacterium tuberculosis (strain H37RV)

C/Species: Mycobacterium tuberculosis

C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000

C/Accession: A70912

R/Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A/Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrett, B.G.

A/Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A/Reference number: A70500; MUID:98295987; PMID:9634230

A/Accession: A70912

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-969 <COL>

A/Cross-references: GB:Z80775; GB:AL123456; NID:93250715; PIDN:CAB02520.1; PID:91568580

A/Experimental source: strain H37RV

A/Genetics:

A/Superfamily: leucine-tRNA ligase

Query Match 7.5%; Score 78.5; DB 2; Length 969;

Best Local Similarity 26.3%; Pred. No. 40;

Matches 36; Conservative 16; Mismatches 68; Indels 17; Gaps 6;

QY 2 GQAKNDATETIPLGEYPEPPELENNKTMRAENG---GRPHHPETKDYSEYSC 57

DB 115 GPDAGLPAEOYAVQGTGHPR--TRTEANVNVFRQLRLGFGHDSRRSFTTDDVDFRW 172

QY 58 RELHFRVYTDGFCRSAPKPVTELVSGGCGPARLLPNAIGKMWKPSGDPDRCPD 114

DB 173 TQWITLQTYNMFDTTAKARPISLVAEPESG--AKCLD---GGRDWAKLTAGEADVID 228

QY 115 R---YRAQRYVOLLCPG 127

DB 229 EYRLVYRADSLVWCPG 245

Search completed: March 28, 2003, 14:19:46

Job time : 29.36 secs

•
•
•
•

GenCore version 5.1.4_p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:11:46 ; Search time 14.6933 Seconds
(without alignments)
536.332 Million cell updates/sec

Title: US-09-867-274-2

Perfect score: 1049
Sequence: 1 OGMQAFKNDATETIPELGEY.....KPRPRASAKNOALEENAY 190

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1049	100.0	213	1	SOST_HUMAN
2	1033	98.5	213	1	SOST_CERAE
3	974	92.9	213	1	SOST_RAT
4	944	90.0	211	1	SOST_MOUSE
5	901.5	85.9	176	1	SOST_BOVIN
6	95.5	9.1	493	1	AR12_HUMAN
7	94.5	9.0	159	1	LSHB_MELGA
8	89.5	8.5	492	1	AR12_MOUSE
9	87	8.3	958	1	HIG_DROME
10	85	8.1	180	1	DAN_HUMAN
11	83.5	8.0	141	1	LSHB_TRIIV
12	81.5	7.8	870	1	BCAL_HUMAN
13	80	7.6	720	1	CTSE_HUMAN
14	80	7.6	1480	1	SLIT_DROME
15	79.5	7.6	318	1	NK2E_MOUSE
16	79	7.5	118	1	LSHB_PHYCA
17	79	7.5	178	1	DAN_RAT
18	79	7.5	355	1	KLF2_HUMAN
19	78.5	7.5	166	1	LSHB_COTJA
20	78.5	7.5	969	1	STL_MYCTU
21	78	7.4	1210	1	EGFR_HUMAN
22	77.5	7.4	951	1	SFR8_HUMAN
23	77	7.3	118	1	LSHB_BALAC
24	77	7.3	510	1	IAL_HUMAN
25	76.5	7.3	220	1	NOL3_MOUSE
26	76.5	7.3	1353	1	SHK2_HUMAN
27	76	7.2	398	1	MUB1_XENLA
28	76	7.2	1213	1	JAG3_BRERA
29	76	7.2	2142	1	BAT2_HUMAN
30	75.5	7.2	318	1	NK2E_RAT
31	75.5	7.2	1436	1	WC11_BOVIN
32	75	7.1	141	1	LSHB_CERST
33	74.5	7.1	972	1	STL_MYCLE

34	74	7.1	138	1	LSHB_MACRU	046483 macropus ru
35	74	7.1	269	1	YAS3_ARCFU	029209 archaeoglob
36	74	7.1	470	1	KLF4_HUMAN	043474 homo sapien
37	74	7.1	474	1	KLF4_MOUSE	060793 mus musculu
38	74	7.1	544	1	H15_DROME	094890 drosophila
39	74	7.1	702	1	HELS_AERPE	094698 aetopyrum p
40	74	7.1	877	1	INCE_CHICK	P53352 gallus gall
41	74	7.1	2715	1	TRX2_HUMAN	Q9urn6 homo sapien
42	73.5	7.0	682	1	SNK_MOUSE	P53351 mus musculu
43	73.5	7.0	682	1	SNK_RAT	Q9r012 rattus norv
44	73.5	7.0	685	1	SNK_HUMAN	Q9ny73 homo sapien
45	73	7.0	178	1	DAN_MOUSE	Q61477 mus musculu

ALIGNMENTS

RESULT 1
SOST_HUMAN STANDARD; PRT; 213 AA.
AC O9BOB4;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sclerostin precursor.
GN SOST.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed:1181578;
RA Bailemans W., Ebeling M., Patel N., van Hul E., Olson P., Dioszegi M.,
RA Laccia C., Wuyts W., van den Ende J., Willems P., Paes-Alves A.F.,
RA Hill S., Bueno M., Ramos F.J., Tacconi P., Dikkers F.G., Stratakis C.,
RA Lindpaintner K., Vickery B., Foerzner D., Van Hul W.;
RT "Increased bone density in sclerosteosis is due to the deficiency of a
RT novel secreted protein (SOST).";
RL Hum. Mol. Genet. 10:537-543(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX PubMed:1179006;
RA Brunkow M.E., Gardner J.C., Van Ness J., Paepers B.W., Kovacevich B.R.,
RA Prohl S., Skonier J.E., Zhao L., Sabo P.J., Fu Y.H., Alisch R.S.,
RA Giller L., Colbert T., Tacconi P., Galas D., Hamersma H.,
RA Beighton P., Mulligan J.T.;
RT "Bone dysplasia sclerosteosis results from loss of the SOST gene
RT product, a novel cystine knot-containing protein.";
RL Am. J. Hum. Genet. 68:577-589(2001).
CC - FUNCTION: Seems to play a role in bone homeostasis.
CC - SUBCELLULAR LOCATION: Secreted (Potential).
CC - TISSUE SPECIFICITY: Widely expressed at low levels with highest
CC levels in bone, cartilage, kidney, liver, bone marrow and primary
CC osteoblasts differentiated for 21 days.
CC - DISEASE: Defects in SOST are the cause of sclerosteosis, a
CC progressive sclerosing bone dysplasia.
CC - SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
CC or send an email to license@ebi.ac.uk).
CC
CC EMBL; AF331844; AAK16158.1; -
CC EMBL; AF326736; AAK13451.1; -
CC EMBL; AF326739; AAK13454.1; -
CC Genbank; HGNC:13771; SOST.
CC MIM; 605740; -
CC MIM; 269500; -

DR InterPro; IPR000359; Cys_knot.
DR SMART; SM00041; CT; 1.
DR PROSITE; PS01185; CTCK_1; FALSE_NEG.
DR PROSITE; PS01225; CTCK_2; FALSE_NEG.
KW Signal; Glycoprotein.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 213 SCLEOSTIN.
FT DOMAIN 82 172 CTCK.
FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 213 AA; 24030 MW; 30DD55CE73D5B82 CRC64;
Query Match 100.0%; Score 1049; DB 1; Length 213;
Best local similarity 100.0%; Pred. No. 5e-88;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 QGWAQKNDATETIIPELGEYEPPEPELENNKTMNRAENGSRPHHPETKDVSEYSCREL 60
DB 24 QGWAQKNDATETIIPELGEYEPPEPELENNKTMNRAENGSRPHHPETKDVSEYSCREL 83
QY 61 HFTRYVTDGPRSAKPYTELVCSCGCGPARLLPNAIGRGMWRSPGDFRCIPDRYRAQR 120
DB 84 HFTRYVTDGPRSAKPYTELVCSCGCGPARLLPNAIGRGMWRSPGDFRCIPDRYRAQR 143
QY 121 VQLCPGGAAPRARKVRLVASCKCKRLTRFHNQSELKDFGEARPOKGRKPRPARSAK 180
DB 144 VQLCPGGAAPRARKVRLVASCKCKRLTRFHNQSELKDFGEARPOKGRKPRPARSAK 203
QY 181 ANQAELENAY 190
DB 204 ANQAELENAY 213
RESULT 2
SOST_CERAE STANDARD; PRT; 213 AA.
ID SOST_CERAE
AC Q9BG78;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sclerostin precursor.
GN SOST.
OS Cercopithecus aethiops (Green monkey) (Grivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21090529; PubMed=11179006;
RA Brunkow M.E., Gardner J.C., Van Ness J., Paepers B.W., Kovacevich B.R.,
RA Proll S., Skonier J.E., Zhao L., Sabo P.J., Fu Y.H., Alish R.S.,
RA Gillett L., Colbert T., Tacconi P., Galas D., Hamerema H.,
RA Beighton P., Mulligan J.T.;
RT "Bone dysplasia sclerosteosis results from loss of the SOST gene
product, a novel cysteine knot-containing protein.";
RL Am. J. Hum. Genet. 68:577-589(2001).
CC -!- FUNCTION: Seems to play a role in bone homeostasis (By
similarity).
CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
CC -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF326742; AAK13457.1; -
CC InterPro; IPR000359; Cys_knot.
CC PROSITE; PS01185; CTCK_1; FALSE_NEG.
CC PROSITE; PS01225; CTCK_2; FALSE_NEG.
CC SMART; SM00041; CT; 1.

DR PROSITE; PS01185; CTCK_1; FALSE_NEG.
DR PROSITE; PS01225; CTCK_2; FALSE_NEG.
KW Signal; Glycoprotein.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 213 SCLEOSTIN.
FT DOMAIN 82 172 CTCK.
FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 213 AA; 23908 MW; 6DA7B5E674728A CRC64;
Query Match 98.5%; Score 1033; DB 1; Length 213;
Best local similarity 98.4%; Pred. No. 1.4e-86;
Matches 187; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 QGWAQKNDATETIIPELGEYEPPEPELENNKTMNRAENGSRPHHPETKDVSEYSCREL 60
DB 24 QGWAQKNDATETIIPELGEYEPPEPELENNKTMNRAENGSRPHHPETKDVSEYSCREL 83
QY 61 HFTRYVTDGPRSAKPYTELVCSCGCGPARLLPNAIGRGMWRSPGDFRCIPDRYRAQR 120
DB 84 HFTRYVTDGPRSAKPYTELVCSCGCGPARLLPNAIGRGMWRSPGDFRCIPDRYRAQR 143
QY 121 VQLCPGGAAPRARKVRLVASCKCKRLTRFHNQSELKDFGEARPOKGRKPRPARSAK 180
DB 144 VQLCPGGAAPRARKVRLVASCKCKRLTRFHNQSELKDFGEARPOKGRKPRPARSAK 203
QY 181 ANQAELENAY 190
DB 204 ANQAELENAY 213
RESULT 3
SOST_RAT STANDARD; PRT; 213 AA.
ID SOST_RAT
AC Q99F67;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sclerostin precursor.
GN SOST.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Sprague-Dawley;
RX MEDLINE=21090529; PubMed=11179006;
RA Brunkow M.E., Gardner J.C., Van Ness J., Paepers B.W., Kovacevich B.R.,
RA Proll S., Skonier J.E., Zhao L., Sabo P.J., Fu Y.H., Alish R.S.,
RA Gillett L., Colbert T., Tacconi P., Galas D., Hamerema H.,
RA Beighton P., Mulligan J.T.;
RT "Bone dysplasia sclerosteosis results from loss of the SOST gene
product, a novel cysteine knot-containing protein.";
RL Am. J. Hum. Genet. 68:577-589(2001).
CC -!- FUNCTION: Seems to play a role in bone homeostasis (By
similarity).
CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
CC -!- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF326741; AAK13456.1; -
CC InterPro; IPR000359; Cys_knot.
CC PROSITE; PS01185; CTCK_1; FALSE_NEG.
CC PROSITE; PS01225; CTCK_2; FALSE_NEG.
CC SMART; SM00041; CT; 1.

FT SIGNAL 1 28 POTENTIAL.
 FT CHAIN 29 213 SCLEROSTIN.
 FT DOMAIN 82 172 CTCK.
 FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 213 AA; 23974 MW; 6C56C878C8D684B CRC64;

Query Match 92.9%; Score 974; DB 1; Length 213;
 Best Local Similarity 92.1%; Pred. No. 3e-81; Mismatches 8; Indels 0; Gaps 0;
 Matches 175; Conservative 7;

QY 1 OGQWAFKNDATETIIPELGEYEPPEPELENNKTNRAENGRRPHHPFERKDVSEYSREL 60
 DB 24 OGQWAFKNDATETIIPGLREYEPPEPELENNQTNRAENGRRPHHPDYTKDVSEYSREL 83
 QY 61 HFTRYVDGPRCSAKPVTLYVCSGCCGPARLLPNAIGRKWMRPSGDPFRCTPDRYRAOR 120
 DB 84 HFTRYVDGPRCSAKPVTLYVCSGCCGPARLLPNAIGRKWMRPSGDPFRCTPDRYRAOR 143
 QY 121 VOLLCGGEAPRARKVRLVASCCKRLTRFHNOSLKDGTENARPQKGRKPRPARSAK 180
 DB 144 VOLLCGGEAPRARKVRLVASCCKRLTRFHNOSLKDGTENARPQKGRKPRPARSAK 203
 QY 181 ANQAELENAY 190
 DB 204 ANQAELENAY 213

RESULT 4
 SOST_MOUSE STANDARD; PRT; 211 AA.
 ID SOST_MOUSE Q99P68; Q99D17;
 AC 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Sclerostin precursor.
 GN SOST.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NC NCB1_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SVJ;
 RX MEDLINE=21090529; PubMed=11179006;
 RA Brunkow M.E., Gardner J.C., Van Ness J., Paepers B.W., Kovacevich B.R.,
 RA Prohl S., Skonier J.E., Zhao L., Sabo P.J., Fu Y.H., Alisch R.S.,
 RA Gillett L., Colbert T., Tacconi P., Galas D., Hamersma H.,
 RA Beighton P., Mulligan J.T.;
 RA "bone dysplasia sclerosteosis results from loss of the SOST gene
 RA product, a novel cystine knot-containing protein";
 RA Am. J. Hum. Genet. 68:577-589(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Head;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi U., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Pleischmann W., Gaasterland T., Glassi C., King B., Kochiwa H.,
 RA Kuell P., Lewis S., Matsumoto Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schiml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okada T., Furuno M., Aono H., Baladrelli R., Baran G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
 RA Guernicchi S., Hill D., Hofmann M., Hume D.A., Kamaya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombert P.,
 RA Norone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Saeki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wymshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohlschütter S.,

RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection";
 RL Nature 409:685-690(2001).
 CC -1- FUNCTION: Seems to play a role in bone homeostasis (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted (Potential).
 CC -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE (CTCK) DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AF326740; AK13455.1; -
 DR EMBL; AK017295; BAB30678.1; -
 DR EMBL; AF326737; AK13452.1; -
 DR MGI; MGI:1921749; Sost.
 DR InterPro; IPR000359; Cys_knot.
 DR PROSITE; PS01185; CTCK_1; FALSE NEG.
 DR PROSITE; PS01225; CTCK_2; FALSE NEG.
 KW Signal; Glycoprotein.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 211 SCLEROSTIN.
 FT DOMAIN 80 170 CTCK.
 FT CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 173 173 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 72 72 G -> D (IN REF. 2).
 SQ SEQUENCE 211 AA; 23443 MW; AEB094E358B34961 CRC64;

Query Match 90.0%; Score 944; DB 1; Length 211;
 Best Local Similarity 88.9%; Pred. No. 1.5e-78; Mismatches 9; Indels 2; Gaps 1;
 Matches 169; Conservative 10;

QY 1 OGQWAFKNDATETIIPELGEYEPPEPELENNKTNRAENGRRPHHPFERKDVSEYSREL 60
 DB 24 OGQWAFKNDATETIIPGLREYEPPEPELENNQTNRAENGRRPHHPDYTKDVSEYSREL 81
 QY 61 HFTRYVDGPRCSAKPVTLYVCSGCCGPARLLPNAIGRKWMRPSGDPFRCTPDRYRAOR 120
 DB 82 HFTRYVDGPRCSAKPVTLYVCSGCCGPARLLPNAIGRKWMRPSGDPFRCTPDRYRAOR 141
 QY 121 VOLLCGGEAPRARKVRLVASCCKRLTRFHNOSLKDGTENARPQKGRKPRPARSAK 180
 DB 142 VOLLCGGEAPRARKVRLVASCCKRLTRFHNOSLKDGTENARPQKGRKPRPARSAK 201
 QY 181 ANQAELENAY 190
 DB 202 ANQAELENAY 211

RESULT 5
 SOST_BOVIN STANDARD; PRT; 176 AA.
 ID SOST_BOVIN Q9BQ79;
 AC 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Sclerostin (Fragment).
 GN SOST.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 NC NCB1_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=21090529; PubMed=11179006;
 RA Brunkow M.E., Gardner J.C., Van Ness J., Paepers B.W., Kovacevich B.R.,
 RA Prohl S., Skonier J.E., Zhao L., Sabo P.J., Fu Y.H., Alisch R.S.,
 RA Gillett L., Colbert T., Tacconi P., Galas D., Hamersma H.,

RA Beighton P., Mulligan J.T.;
 RT "bone dysplasia cysteroestosis results from loss of the SOST gene
 RT product a novel cysteine knot-containing protein.";
 RL Am. J. Hum. Genet. 68:577-589(2001).
 CC -1- FUNCTION: Seems to play a role in bone homeostasis (By
 CC similarity).
 CC -1- SUBCELLULAR LOCATION: Secreted (Potential).
 CC -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTEINE KNOT-LIKE (CTCK) DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF326738; AAK13453.1; -;
 CC InterPro; IPR000359; Cys_knot.
 CC SMART; SM00041; CT; 1.
 CC PROSITE; PS01185; CTCK_1; FALSE NEG.
 CC PROSITE; PS01225; CTCK_2; FALSE NEG.
 CC Glycoprotein.
 FT NON TER 1 1
 FT DOMAIN 51 141 CTCK.
 FT CARBOHYD 22 22 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 144 144 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT NON TER 176 176
 SQ SEQUENCE 176 AA; 19743 MW; 35F0CA61A425F4DB CRC64;
 Query Match 85.9%; Score 901.5; DB 1; Length 176;
 Best Local Similarity 93.8%; Pred. No. 8.9e-75;
 Matches 166; Conservative 2; Mismatches 8; Indels 1; Gaps 1;
 QY 8 NDATETIIPELGGEYEPPELENNKTKMRAENGRRPHHPETKDVSYSCRELIHFRYVT 67
 DB 1 NDATETIIPELGGEYEPPELENNKTKMRAENGRRPHHPETKDVSYSCRELIHFRYVT 59
 QY 68 DGPGRSAKPVTELVCSGCGCPARLLPNAIGRKWMPSPSGDFCIPRYVAQSVOLLCPG 127
 DB 60 DGPGRSAKPVTELVCSGCGCPARLLPNAIGRKWMPSPSGDFCIPRYVAQSVOLLCPG 119
 QY 128 GEAPRRARKVRLVASCCKRLTRFRNOSGLKDFGTEARPOKGRKPRPRASAKANQA 184
 DB 120 GAAPRRARKVRLVASCCKRLTRFRNOSGLKDFGTEARPOKGRKPRPRASAKANQA 176
 RESULT 6
 ARI2 HUMAN STANDARD; PRT; 493 AA.
 ID ARI2 HUMAN STANDARD; PRT; 493 AA.
 AC 095376; 09UEM9; 09HB26;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Ariadne-2 protein homolog (ARI-2) (Triadi protein) (HT005).
 GN ARI2 OR ARI2 OR TRIAD1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 OX 1
 RN SEQUENCE FROM N.A.
 RP MEDLINE=9349709; Pubmed=10422847;
 RA van der Reijden B.A., Erpelinck-Verschueren C.A.J., Loewenberg B.,
 RA Jansen J.H.;
 RT "Riids: a new class of proteins with a novel cysteine-rich
 RT signature";
 RL Protein Sci. 8:1557-1561(1999).
 RN 12
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20341325; Pubmed=10860484;
 RA Aguilera M., Oliveros M., Martinez-Padon M., Barbas J.A., Ferrus A.;
 RA "Ariadne-1: a vital Drosophila gene is required in development and

RT defines a new conserved family of ring-finger proteins.";
 RL Genetics 155:1231-1244(2000).
 RN 13
 RP SEQUENCE FROM N.A.
 RC TISUB=Hypothalamus;
 RX MEDLINE=20402571; Pubmed=10931946;
 RA Hu R.-M., Han Z.-G., Song H.-D., Peng Y.-D., Huang Q.-H., Ren S.-X.,
 RA Gu Y.-Y., Huang C.-H., Li Y.-B., Jiang C.-L., Fu G., Zhang Q.-H.,
 RA Gu B.-W., Dai M., Mao Y.-F., Gao G.-F., Rong R., Ye M., Zhou J.,
 RA Xu S.-H., Gu J., Shi J.-X., Jin W.-R., Zhang C.-K., Wu T.-M.,
 RA Huang G.-Y., Chen Z., Chen M.-D., Chen J.-L.;
 RT "Gene expression profiling in the human hypothalamus-pituitary-adrenal
 RT axis and full-length cDNA cloning.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:9543-9548(2000).
 RN 14
 RP SEQUENCE FROM N.A.
 RC TISUB=Muscle;
 RA Strausberg R.;
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: MIGHT ACT AS AN E3 UBIQUITIN-PROTEIN LIGASE, OR AS PART
 CC OF E3 COMPLEX, WHICH ACCEPTS UBIQUITIN FROM SPECIFIC E2 UBIQUITIN-
 CC CONJUGATING ENZYMES, SUCH AS UBE213/UBC4, AND THEN TRANSFERS IT
 CC TO SUBSTRATES.
 CC -1- SUBUNIT: INTERACTS WITH UBE213 (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- SIMILARITY: CONTAINS 2 RING-TYPE ZINC FINGERS.
 CC -1- SIMILARITY: CONTAINS 1 IRR-TYPE ZINC FINGER.
 CC -1- CAUTION: REP.3 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO
 CC FRAMESHIFTS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF099149; AAC82469.1; -;
 CC EMBL; AJ130978; CAA10276.1; -;
 CC EMBL; AF183427; AAG09696.1; ALT_FRAME.
 CC EMBL; BC000422; AAH00422.1; -;
 CC GeneW; HGNC:690; ARI2.
 DR MIM; 605615; -;
 DR InterPro; IPR002867; Znf_GCHC.
 DR InterPro; IPR001879; Znf_CCHC.
 DR InterPro; IPR001841; Znf_Fing.
 DR Pfam; PF01485; IRR; 2.
 DR SMART; SM0184; RING; 2.
 DR SMART; SM00343; Znf_C2HC; 1.
 DR PROSITE; PS00518; ZF_RING_1; 1.
 DR PROSITE; PS50089; ZF_RING_2; 2.
 KW ubl conjugation pathway; Nuclear protein; Coiled coil; Zinc-finger;
 KW Repeat.
 FT DOMAIN 4 75 ASP/GLU-RICH (ACIDIC).
 FT DOMAIN 22 29 POLY-GLU.
 FT ZN_FING 139 192 RING-TYPE 1.
 FT ZN_FING 208 270 IRR-TYPE.
 FT ZN_FING 300 344 RING-TYPE 2.
 FT DOMAIN 369 400 COILED COIL (POTENTIAL).
 FT DOMAIN 439 492 COILED COIL (POTENTIAL).
 FT CONFLICT 280 281 CA -> LQ (IN REF. 2).
 SQ SEQUENCE 493 AA; 57818 MW; 30AFPD327B51013 CRC64;
 Query Match 9.1%; Score 95.5; DB 1; Length 493;
 Best Local Similarity 24.2%; Pred. No. 0.2;
 Matches 45; Conservative 24; Mismatches 54; Indels 63; Gaps 11;
 QY 3 WQA-----FKNDATETIIPELGGEYEPPELENNKTKMRAENGRRPHHPH-----PFETKD 51
 DB 100 WQVSEILIRYKSNSSQQLVERVOQNPESHVPTS-----HPPHCAVCMQFVRKE 149
 QY 52 -----VSEY-----SCRELIHFRYVTDGPRSAKPVTELVCSGCGCPAR-----LLPN 94

```

Db 150 NLLSLACOHQFCRSCWBOHOSVJXKG-----VGVGVSCMAQDCFLRPEDFVPLLPN 203
Qy 95 AIGRGKMMRSGDFRCPDRYRAORVOLLCPGGEA-----PRARKVRL-----VAS 141
Db 204 BELREKRRRLTFRDY--VESHYOLQ-----LCPADCPMTIRVOEPARRVQCRCNEVFC 257
Qy 142 CKCKRL 147
Db 258 FKCRQM 263

RESULT 7
LSHB_MELGA STANDARD; PRT; 159 AA.
ID ID LSHB_MELGA STANDARD; PRT; 159 AA.
AC P45646;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lutropin beta chain precursor (Luteinizing hormone beta subunit) (LSH-
beta) (LSH-B) (LH-B).
GN LHB.
OS Meleagris gallopavo (Common turkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Meleagrididae; Meleagris.
OX NCBI_TaxId=9103;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary;
RX MEDLINE=95290073; PubMed=7772235;
RA You S., Foster L.K., Silsby J.L., el Halawani M.E., Foster D.N.;
RT "Sequence analysis of the turkey LH beta subunit and its regulation
RT by gonadotrophin-releasing hormone and prolactin in cultured
RT pituitary cells."
RL J. Mol. Endocrinol. 14:117-129(1995).
RN [2]
CC -1- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING
CC THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.
CC -1- SUBUNIT: HETERODIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
CC CHAIN WHICH CONFERES BIOLOGICAL SPECIFICITY TO THYROTROPIN,
CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
CC -1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, L35519; AAA74125.1; ALT_INIT.
DR HSSP; P01233; 1XUL.
DR InterPro; IPR000359; Cys_knot.
DR InterPro; IPR002400; GF_cysknot.
DR InterPro; IPR001545; Gly_hormoneB.
DR Pfam; PF00007; Cys_knot_1.
DR PRINTS; PR00438; GFCYSKNOT.
DR SMART; SM00068; GH; 1.
DR PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.
DR PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.
DR Hormone; Signal; Glycoprotein.
DR SIGNAL 1 39
FT CHAIN 40 159 LUTROPIN BETA CHAIN.
FT DISULFID 48 96 BY SIMILARITY.
FT DISULFID 62 111 BY SIMILARITY.
FT DISULFID 65 149 BY SIMILARITY.
FT DISULFID 73 127 BY SIMILARITY.
FT DISULFID 77 129 BY SIMILARITY.
FT DISULFID 132 139 BY SIMILARITY.
FT CARBOHYD 52 52 N-LINKED (GLCNAC... ) (POTENTIAL).
SQ SEQUENCE 159 AA; 16285 MW; 52B50C8C879653C6 CRC64;

```

```

Query Match 9.0%; Score 94.5; DB 1; Length 159;
Best Local Similarity 29.4%; Pred. No. 0.077;
Matches 32; Conservative 7; Mismatches 49; Indels 21; Gaps 5;

Qy 39 GGRPHHPEFKDVSSEYSCRELHFFRYTIDGPCRSAKVTELVSCGGCPARLPMIGR 98
Db 43 GGRPP-----CRPTNVAAVEKDECPQCMATVTTACGCGYCTR---EPVYR 85
Qy 99 GKWRPSPGDFRCPDRYRAORVOLL-CPGGEAPRRARKVRLVASCCKR 146
Db 86 SPLGRP--PQSSCTYGALRYERKALMGCPIDSPRV-LLPVALSRCAR 131

RESULT 8
AR12_MOUSE STANDARD; PRT; 492 AA.
ID ID AR12_MOUSE STANDARD; PRT; 492 AA.
AC Q921K6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ariadne-2 protein homolog (Ari-2) (Triad1 protein) (UbcM4-interacting
DE protein 48).
GN AR12 OR AR12 OR TRIAD1 OR UTP48.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BAUB/C; TISSUE=Liver;
RX MEDLINE=99358765; PubMed=10431818;
RA Martinez-Noel G., Niedenthal R., Tamura T., Harbers K.;
RT "A family of structurally related RING finger proteins interacts
RT specifically with the ubiquitin-conjugating enzyme UbcM4."
RL FEBS Lett. 454:257-261(1999).
RN [2]
CC SEQUENCE FROM N.A.
CC TISSUE=Brain, and Embryo;
CC MEDLINE=20341325; PubMed=10880484;
CC Aguilera M., Oliveros M., Martinez-Padron M., Barbás J.A., Ferrus A.;
RT "Ariadne-1: a vital Drosophila gene is required in development and
RT defines a new conserved family of ring-finger proteins."
RL Genetics 155:1231-1244(2000).
CC -1- FUNCTION: MIGHT ACT AS AN E3 UBIQUITIN-PROTEIN LIGASE, OR AS PART
CC OF E3 COMPLEX, WHICH ACCEPTS UBIQUITIN FROM SPECIFIC E2 UBIQUITIN-
CC CONJUGATING ENZYMES, SUCH AS UBE2L3/UBC4, AND THEN TRANSFERS IT
CC TO SUBSTRATES.
CC -1- SUBUNIT: INTERACTS WITH UBE2L3.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: CONTAINS 2 RING-TYPE ZINC FINGERS.
CC -1- SIMILARITY: CONTAINS 1 IRR-TYPE ZINC FINGER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, AF124664; AAD24573.1; -
DR EMBL, AJ130975; CAA10273.1; -
DR MGD; MGI:1344361; Arih2.
DR InterPro; IPR002867; Znf_CGHC.
DR InterPro; IPR001878; Znf_CGHC.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF01485; IRR; 2.
DR SMART; SM00184; RING; 2.
DR SMART; SM00343; ZNF_C2HC; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00518; ZF_RING_2; 2.
DR Ubl conjugation pathway; Nuclear protein; Coiled coil; Zinc-finger;
KW Repeat.

```

FT DOMAIN 4 74 ASP/GLU-RICH (ACIDIC).
 FT DOMAIN 22 29 POLY-GLU.
 FT ZN FING 138 191 RING-TYPE 1.
 FT DOMAIN 207 269 IIR-TYPE.
 FT ZN FING 299 343 RING-TYPE 2.
 FT DOMAIN 368 399 COILED COIL (POTENTIAL).
 FT DOMAIN 438 491 COILED COIL (POTENTIAL).
 SQ SEQUENCE 492 AA; 57696 MW; B55EA54FEBC3ADAF CRC64;
 Query Match 8.5%; Score 89.5; DB 1; Length 492;
 Best Local Similarity 23.1%; Pred. No. 0.7; Indels 63; Gaps 11;
 Matches 43; Conservative 23; Mismatches 55;
 QY 3 WQA-----FKNDATETIIPELGEPPEPELENNKTMRAENGSRPPHH-----PFETKD 51
 DB 99 WQVEIILDRYNSNAQQLVEARVQNPSSKIVPT-----AHPHHCACVQFVKE 148
 QY 52 -----VSEY-----SCRELHFRVYTDGPCRSKAFVTELVCSGCCGPAP-----LLEN 94
 DB 149 NLLSLACOHCFRCSCWEQHCSVLVKDG-----VGVGISCAQDCPFRTEDEFFVPLPN 202
 QY 95 AIGRKMWRPSPDPFRCPDRYRAORVQLCPGGEA-----PRARKYRL-----VAS 141
 DB 203 EELDKTRRYFRDY--VESHFQLQ-----LCPGADCPMVRVQEPFRARRVQCNKRCSEVFC 236
 QY 142 CKCRKL 147
 DB 257 FKCRQM 262
 RESULT 9
 HIG DROME STANDARD; PRT; 958 AA.
 ID 009101;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Locomotion-related protein Hikaru genki precursor.
 GN HIG.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryote; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OX NBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Head;
 RX MEDLINE=93213498; Pubmed=8461133;
 RA Hoshino M., Matsuzaki F., Nabeshima Y.-I., Hama C.;
 RT "Hikaru genki, a CNS-specific gene identified by abnormal locomotion
 in Drosophila, encodes a novel type of protein.";
 RL Neuron 10:395-407(1993).
 CC -1- FUNCTION: HAS A ROLE IN THE DEVELOPMENT OF CNS FUNCTIONS INVOLVED
 IN LOCOMOTOR ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- ALTERNATIVE PRODUCTS: 4 ISOFORMS; 1, 2, 3 (SHOWN HERE) AND 4; ARE
 PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN PCC NEURONS AND NEUROBLASTS
 IN THE PROCEPHALIC NEUROGENIC REGION IN THE CENTRAL NERVOUS
 SYSTEM.
 CC -1- DEVELOPMENTAL STAGE: MOST ABUNDANT DURING AND/OR AFTER NEURONAL
 DIFFERENTIATION AND DURING CELL SPECIFICATION OR APOGENESIS.
 CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY.
 CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 4 SUSHI (SCR) DOMAINS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; D13884; BAA02984.1; -
 DR EMBL; D13885; BAA02985.1; -
 DR EMBL; D13886; BAA02986.1; -
 DR EMBL; D13887; BAA02987.1; -
 DR HSP; P10998; IVD.
 DR FLYBASE; FBgn010114; hig.
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR003006; IG_MHC.
 DR InterPro; IPR000436; Sushi_SCR_CCP.
 DR Pfam; PF00047; Ig_1.
 DR Pfam; PF00084; sushi; 5.
 DR SMART; SM00032; CCP; 5.
 DR SMART; SM00409; IG; 1.
 DR Glycoprotein; Alternative splicing; Immunoglobulin domain; Repeat;
 KW Signal, signal.
 FT SIGNAL 1 31 POTENTIAL.
 FT CHAIN 32 958 LOCOMOTION-RELATED PROTEIN HIKARU GENKI.
 FT DOMAIN 630 709 IG-LIKE C2-TYPE DOMAIN.
 FT DOMAIN 713 769 SUSHI 1.
 FT DOMAIN 772 828 SUSHI 2.
 FT DOMAIN 831 891 SUSHI 3.
 FT DOMAIN 893 952 SUSHI 4.
 FT SITE 318 320 CELL ATTACHMENT SITE.
 FT DISULFID 714 755 BY SIMILARITY.
 FT DISULFID 741 768 BY SIMILARITY.
 FT DISULFID 773 814 BY SIMILARITY.
 FT DISULFID 800 827 BY SIMILARITY.
 FT DISULFID 832 877 BY SIMILARITY.
 FT DISULFID 863 890 BY SIMILARITY.
 FT DISULFID 894 939 BY SIMILARITY.
 FT DISULFID 922 952 BY SIMILARITY.
 FT CARBOHYD 376 376 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 525 525 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 605 605 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 620 620 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 752 752 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 789 789 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 529 553 MISSING (IN ISOFORM 2 AND ISOFORM 4).
 FT VARSPLIC 892 958 MISSING (IN ISOFORM 1 AND ISOFORM 2).
 SQ SEQUENCE 958 AA; 107027 MW; 4161258B85ABC764 CRC64;
 Query Match 8.3%; Score 87; DB 1; Length 958;
 Best Local Similarity 19.7%; Pred. No. 2.3;
 Matches 47; Conservative 24; Mismatches 82; Indels 86; Gaps 8;
 QY 6 FKNDATETIIPELGEPPEPELENN----- 30
 DB 343 FKGDGOLPPESGIGPEPEPLADQLQYGNQSSARVALMQRKSGRTAGALS 402
 QY 31 -----KTMNRAENG-----GRPHHPETK-DVSEYSCRELHFRVYTDGPCRS 73
 DB 403 RKGGGDDSSKTSKXKGIYDEERAGYTHPDDPEDEEEDVDIIOQTFE----- 457
 QY 74 AKPTVELVCSGCCGP-----ARLLPNAIGRKMWRPSPDPFRCPDRYRAORVQLCPGGEA 130
 DB 458 -----VSEIIFPGEIIPGMGBRCLKIRCVAKWVG-----LCATNE 495
 QY 131 PRARKYR--LVASCKCRKLTRFNQSELDKFGTEARPPQKGRPRPARAKANQVEL 186
 DB 496 DDNGNVKQPIYKSKCHVNRIPS-HLLSYRNIISVTPIPPNMGWKRTRLSKSTLLNSTEI 553
 RESULT 10
 ID DAN_HUMAN STANDARD; PRT; 180 AA.
 AC P41271;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neuroblastoma suppressor of tumorigenicity 1 (Zinc finger protein
 DAN) (N03).
 GN NBI OR DAN.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 RN NCB1_TaxID=9606;
 RP [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=9436724; PubMed=8084583;
 RA Enomoto H., Ozaki T., Takahashi E., Nomura N., Tabata S.,
 RA Takahashi H., Ohnuma N., Tanabe M., Iwai J., Yoshida J., Matsunaga T.,
 RA Sakiyama S.;
 RT "Identification of human DAN gene, mapping to the putative
 RT neuroblastoma tumor suppressor locus.";
 RL Oncogene 9:2785-2791(1994).
 CC -1- FUNCTION: POSSIBLE CANDIDATE AS A TUMOR SUPPRESSOR GENE OF
 CC NEUROBLASTOMA. MAY PLAY AN IMPORTANT ROLE IN PREVENTING CELLS
 CC FROM ENTERING THE FINAL STAGE (G1/S) OF THE TRANSFORMATION
 CC PROCESS.
 CC -1- TISSUE SPECIFICITY: MOST ABUNDANT IN NORMAL LUNG AND MENINGIOMA.
 CC -1- DISEASE: THE LOSS OF DAN GENE BY MUTATION IS POSSIBLY THE CAUSE OF
 CC THE DEVELOPMENT AND/OR PROGRESSION OF HUMAN NEUROBLASTOMA.
 CC -1- SIMILARITY: HIGH, TO OTHER MAMMALIAN DAN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D28124; BAA05671.1; -;
 DR Genew; HGNC:7650; NBL1.
 DR MIM: 600613; -;
 DR InterPro; IPR000359; Cys knot.
 DR InterPro; IPR004133; DAN_dom.
 DR Pfam; PF03045; DAN; 1.
 DR SMART; SM00041; CT; 1.
 DR Anti-oncogene; DNA-binding; Zinc-finger.
 DR ZN FING 81 101 CA-TYPE (POTENTIAL).
 FT DOMAIN 141 171 PRO-RICH
 FT SEQUENCE 180 AA; 19277 MW; 15233229DB65865 CRC64;
 SQ
 Query Match 8.1%; Score 85; DB 1; Length 180;
 Best Local Similarity 26.0%; Pred. No. 0.63;
 Matches 33; Conservative 15; Mismatches 61; Indels 18; Gaps 7;
 QY 57 CRELHTRVYTDGPGCRSAKPVTELVCSGGCGPARLLPNAIGRGKMRPSPGPDF-----RC 111
 DB 34 CEAKNITQIVGHSQC-EAKSIQNRACIGQCF-SYVYNTF-----POSTESLVHCDSC 84
 QY 112 IPRVYAGRVOLLCPG-GEAPRARK-VRLVASCKKRLRFHNSQLXCF--GTEARQ 167
 DB 85 MPAGSMWEIVTLECPHBEVPRVDKLVKILHSCQACGKPSHESLYVVGEDGPGSQ 144
 QY 168 KGRKRP 174
 DB 145 PGTNPR 151
 RESULT 11
 LSHB TRIUV STANDARD; PRT; 141 AA.
 ID LSHB TRIUV
 AC 046482;
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE Lutropin beta chain precursor (luteinizing hormone beta subunit) (LSH-
 DE beta) (LSH-B) (LH-B).
 GN LHB.
 OS Trichosurus vulpecula (Brush-tailed possum).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Metatheria; Diprotodontia; Phalangeridae; Trichosurus.

OK NCB1_TaxID=9337;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Plutary;
 RX MEDLINE=9834524; PubMed=9680384;
 RA Harrison G.A., Deane E.M., Cooper D.W.;
 RT "cDNA cloning of luteinizing hormone subunits from brushtail possum
 RT and red kangaroo.";
 RL Mamm. Genome 9:638-642(1998).
 RP [2]
 RP SEQUENCE FROM N.A.
 RA Lawrence S.B., McNatty K.P., Fidler A.E.;
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING
 CC THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.
 CC -1- SUBUNIT: HETEROOLIGOMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA
 CC CHAIN WHICH CONFERES BIOLOGICAL SPECIFICITY TO THYROTROPIN,
 CC LUTROPIN, FOLLITROPIN AND GONADOTROPIN.
 CC -1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF017448; AAC96019.1; -;
 DR EMBL; AF090388; AAC63526.1; -;
 DR HSSP; P01233; IXUL.
 DR InterPro; IPR000359; Cys knot.
 DR InterPro; IPR001545; Gly hormone.
 DR Pfam; PF00007; Cys knot; 1.
 DR SMART; SM00068; GHE; 1.
 DR PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.
 DR PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.
 DR Hormone; Signal; Glycoprotein.
 FT SIGNAL 1 22
 FT CHAIN 23 141 LUTROPIN BETA CHAIN.
 FT DISULFID 30 78 BY SIMILARITY.
 FT DISULFID 44 93 BY SIMILARITY.
 FT DISULFID 47 93 BY SIMILARITY.
 FT DISULFID 55 109 BY SIMILARITY.
 FT DISULFID 59 111 BY SIMILARITY.
 FT DISULFID 114 121 BY SIMILARITY.
 FT CARBOHYD 34 34 N-LINKED (GLCNAC...) (POTENTIAL).
 FT SEQUENCE 141 AA; 15060 MW; C6CF98036B3C4EB0 CRC64;
 SQ
 Query Match 8.0%; Score 83.5; DB 1; Length 141;
 Best Local Similarity 27.7%; Pred. No. 0.67;
 Matches 33; Conservative 10; Mismatches 59; Indels 17; Gaps 5;
 QY 57 CRELHTRVYTDGPGCRSAKPVTELVCSGGCGPARLLPNAIGRGKMRPSPGPDFRCIPDR 115
 DB 30 CRPTNATLAESDACPVCVTFITICAGYCPSMVRLPAL-----PSPQVLTYRE 82
 QY 116 YRAQRVOLL-CPGGEAP-RARKVRLVASCKKRLTRFNSQLXCF--GTEARQ 172
 DB 83 LSFSSIRLPGCPGVDPITFSFVALSCGSCGRSLH-----SDCGPRPRLCTRP 134
 RESULT 12
 BCAL HUMAN STANDARD; PRT; 870 AA.
 ID BCAL HUMAN
 AC P56945;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE CRK-associated substrate (p130Cas) (Breast cancer anti-estrogen
 DE resistance 1 protein).
 GN BCAR1 OR CRKAS OR CAS.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 CC NCBI_TaxID=9606;
 RN
 RP SEQUENCE FROM N.A.
 RC TISSUE=Breast cancer;
 RX MEDLINE=20119429; PubMed=10639512;
 RA Brinkman A., van der Flier S., Kok E.M., Dorasari L.C.J.;
 RT "BCAR1, a human homologue of the adapter protein p130Cas, induces
 RT anti-estrogen resistance in breast cancer cells.";
 RL J. Natl. Cancer Inst. 92:112-120(2000).
 RN
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA Otto E., Birnbaum S., Verbeek M., Hildebrandt F.;
 RT "Interaction between human Crk-associated substrate (p130Cas) and
 RT nephrin.";
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
 RN
 RP SEQUENCE FROM N.A.
 RC TISSUE=Cornea;
 RA Imoto Y., Ohguro N., Yoshida A., Tsujikawa M., Inoue Y., Tano Y.;
 RT "The effects of growth factors on tyrosine phosphorylation of p130Cas
 RT in corneal epithelial cell.";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
 RA Nishikawa T., Nagai K., Sugano S., Aotsuma S., Yoshikawa Y.,
 RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,
 RA Nakamura Y., Nagahara K., Masuko Y., Sasaki N.;
 RT "NEBO human cDNA sequencing project.";
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 CC
 CC -1- FUNCTION: DOCKING PROTEIN WHICH PLAYS A CENTRAL COORDINATING ROLE
 CC FOR TYROSINE-KINASE-BASED SIGNALING RELATED TO CELL ADHESION.
 CC IMPLICATED IN INDUCTION OF CELL MIGRATION, OVEREXPRESSION CONFERS
 CC ANTITUMOR RESISTANCE ON BREAST CANCER CELLS.
 CC
 CC -1- SUBUNIT: FORMS COMPLEXES IN VIVO WITH FOCAL ADHESION KINASE 1,
 CC ADAPTER PROTEIN CRKL AND LYN KINASE. CAN HETERODIMERIZE WITH CASL
 CC INTERACTS WITH NEPHROCYSTIN AND PTK2B (BY SIMILARITY).
 CC
 CC -1- SUBCELLULAR LOCATION: FOCAL ADHESIONS. UNPHOSPHORYLATED FORM
 CC LOCALIZES IN THE CYTOPLASM AND CAN MOVE TO THE MEMBRANE UPON
 CC TYROSINE PHOSPHORYLATION (BY SIMILARITY).
 CC
 CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED WITH AN ABUNDANT EXPRESSION
 CC IN THE TESTIS. LOW LEVEL OF EXPRESSION SEEN IN THE LIVER, THYMUS,
 CC AND PERIPHERAL BLOOD LEUCOCYTES. THE PROTEIN HAS BEEN DETECTED IN
 CC A B-CELL LINE.
 CC
 CC -1- DOMAIN: CONTAINS A CENTRAL DOMAIN (SUBSTRATE DOMAIN) CONTAINING
 CC MULTIPLE POTENTIAL SH2-BINDING SITES AND A C-TERMINAL DOMAIN
 CC CONTAINING A DIVERGENT HELIX-LOOP-HELIX (HLH) MOTIF. THE SH2-
 CC BINDING SITES PUTATIVELY BIND CRK, NCK AND ABL SH2 DOMAINS.
 CC THE HLH MOTIF IS ABSOLUTELY REQUIRED FOR THE INDUCTION OF
 CC PSEUDOPHYAL GROWTH IN YEAST AND MEDIATES HETERODIMERIZATION
 CC WITH CASL (BY SIMILARITY).
 CC
 CC -1- DOMAIN: A SERINE-RICH REGION PROMOTES ACTIVATION OF THE SERUM
 CC RESPONSE ELEMENT (SRE).
 CC
 CC -1- DOMAIN: THE SH3 DOMAIN IS NECESSARY FOR THE LOCALIZATION OF THE
 CC PROTEIN TO FOCAL ADHESIONS AND INTERACTS WITH ONE PROLINE-RICH
 CC REGION OF FOCAL ADHESION KINASE 1.
 CC
 CC -1- PTM: FOCAL ADHESION KINASE 1 PHOSPHORYLATES THE PROTEIN AT THE
 CC YDYVHL MOTIF. THE SRC FAMILY KINASES ARE RECRUITED TO THE
 CC PHOSPHORYLATED SITES AND CAN PHOSPHORYLATE OTHER TYROSINE
 CC RESIDUES. TYROSINE PHOSPHORYLATION IS TRIGGERED BY INTEGRIN
 CC MEDIATED ADHESION OF CELLS TO THE EXTRACELLULAR MATRIX.
 CC
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC
 CC -1- SIMILARITY: BELONGS TO THE CAS FAMILY.
 CC
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 CC
 CC EMBL: AJ242987; CAB75875.2; -
 CC EMBL: AF218451; AAF27527.1; -
 CC EMBL: AB040024; BAA92711.1; -
 CC EMBL: AK027608; BAB55230.1; -
 CC HSSP: P07751; 1BK2.
 CC Genew: HGNC:971; BCAR1.
 CC MIM: 602941; -
 CC InterPro: IPR001452; SH3.
 CC Pfam: PF00018; SH3; 1.
 CC PRINTS: PR00452; SH3DOMAIN.
 CC ProDom: PD000066; SH3; 1.
 CC SMART: SM00326; SH3; 1.
 CC PROSITE: PS00002; SH3; 1.
 CC KW Phosphorylation; SH3 domain; SH3-binding; Cell adhesion.
 CC
 CC FT DOMAIN 3 65
 CC FT DOMAIN 74 87
 CC FT DOMAIN 115 416
 CC FT DOMAIN 422 614
 CC FT SITE 635 643
 CC FT DOMAIN 746 796
 CC FT CONFLICT 1 4
 CC FT CONFLICT 63 63
 CC FT CONFLICT 236 236
 CC FT CONFLICT 349 349
 CC FT CONFLICT 363 363
 CC FT CONFLICT 428 428
 CC FT CONFLICT 471 471
 CC SQ SEQUENCE 870 AA; 93361 MW; D57A9C4FCACAFAD0 CRC64;
 CC
 CC Query Match 7.8%; Score 81.5; DB 1; Length 870;
 CC Best Local Similarity 20.9%; Pred. No. 6.7;
 CC Matches 41; Conservative 27; Mismatches 71; Indels 57; Gaps 9;
 CC
 CC QY 9 DATEIPE-----LGEYPPPELENKTMNAENGRRPHHPFTKDVSEYSCRELH 61
 CC DB 112 DSVYLVPTPSAQQGLYVPGSPQSPQSPAKOTSTFSKOTPHHPSPATDLYQ----- 166
 CC QY 62 PTRVYTDGCRSAKPEVTELVCSGQCG-----PAR-LIPNAIGRK 100
 CC DB 167 -----VPPGGGPAQDIYVPPSAGSHDIYVPPSMDSRWSGKTPAKVVPTRVGGY 222
 CC QY 101 WRRSGDPKPCPRIRYKQVQLCPGG-----EAPRAKVLVASCKKLTTRHNGSEL 156
 CC DB 223 VYEAQPE-----QDEYDIPR-HLAPQDIDYVPPVPGT-----LPQYGO-EV 266
 CC QY 157 KDFGTEARPKGRKP 172
 CC DB 267 YTPPMAYKGPNGRDP 282
 CC
 CC RESULT 13
 CC GENE HUMAN
 CC ID GENE HUMAN STANDARD; PRT; 720 AA.
 CC AC Q9NYZ3; Q9BRE0; Q9Y577; Q9UGZ9;
 CC DT 15-JUN-2002 (Rel. 41, Created)
 CC DT 15-JUN-2002 (Rel. 41, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE G2 and S phase expressed protein 1 (B99 homolog).
 CC GN GTSB1.
 CC OS Homo sapiens (Human).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 CC NCBI_TaxID=9606;
 CC
 CC [1]
 CC RN
 CC RP SEQUENCE FROM N.A.
 CC RC TISSUE=Placenta;
 CC RX MEDLINE=20432105; PubMed=10974554;
 CC RA Monte M., Collavin L., Lazarevic D., Uterera R., Dragani T.A.,
 CC Schneider C.;

"Cloning, chromosome mapping and functional characterization of a human homologue of murine Gtse-1 (B99) gene.";
 RT Gene 254:229-236(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20057165; PubMed=10591208;
 RA Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Baare D.M.,
 RA Clump M., Smith L.J., Ainscough R., Almeida J.P., Babajani A.K.,
 RA Bagley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
 RA Bird C.P., Blakey S.E., Bridgman A.M., Buck D., Burgess J.,
 RA Burdill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
 RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
 RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
 RA Dhanu P.D., Dockree C., Dodsworth S.D., Durbin R.M., Ellington A.G.,
 RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,
 RA Gilbert J.G.R., Goward M.E., Graham D.V., Griffiths M.N.D., Hall C.,
 RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,
 RA Hunt S.E., Jones M.C., Kerhaw J., Kimberley A.M., King A.,
 RA Laid G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,
 RA Marlyn I.D., Mashreghi-Mohammadi M., Matthews L.H., McCann O.T.,
 RA Mcclaren S., McMuray A.A., Milne S.A., Morrison B.J.,
 RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Philimore B.J.C.T.,
 RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,
 RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
 RA Soderlund C., Spragon L., Steward C.A., Suleston J.E., Swann R.M.,
 RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
 RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
 RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
 RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
 RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,
 RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
 RA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,
 RA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,
 RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shauli S., Sloan D., Song L.,
 RA Wang Q., Wang Y., Wang Z., White J., Williamson D., Wu H., Yao Z.,
 RA Zhan M., Zhang G., Chisoe S., Murray J., Miller M., Mink P.,
 RA Fulan R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
 RA Cordes M., Du Z., Fullon L., Goela D., Graves T., Hawkins J.,
 RA Hinds K., Kemp K., Latreille P., Layman D., Ozeresky P., Rohlfing T.,
 RA Scheet P., Walker C., Wamsley A., Woldmann P., Pepin K., Nelson J.,
 RA Korn I., Bedell J.A., Hillier L., Mardis E., Waterston R., Wilson R.,
 RA Emanuel B.S., Shaikh T., Kurahashi H., Salter S., Budarf M.L.,
 RA McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E., Edelmann L.,
 RA Kim U.J., Shinya H., Simon M.I., Dumanek J.P., Payard M., Kedra D.,
 RA Serousi E., Franssen I., Tapia I., Bruder C.E., O'Brien K.P.,
 RA Wilkinson P., Bodenteich A., Hartman K., Hu X., Khan A.S., Lane L.,
 RA Tikhunov Y., Wright H.,
 RT "The DNA sequence of human chromosome 22.";
 RL Nature 402:489-495(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Muscle;
 RA Strausberg R.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: May be involved in p53-induced cell cycle arrest in G2/M
 CC phase by interfering with microtubule rearrangements that are
 CC required to enter mitosis. Overexpression delays G2/M phase
 CC progression.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic. Associated with microtubules.
 CC -1- DEVELOPMENTAL STAGE: Expressed in G2/M phase. Not detected in
 CC quiescent cells.
 CC -1- PTM: Phosphorylated in mitosis (By similarity).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF223408; AAF31459.1; -;
 CC EMBL; AL031588; CAB38415.1; -;

DR EMBL; AL022325; CAB63079.1; -;
 DR EMBL; BC006325; AAH06325.1; -;
 KM Microtubules; Phosphorylation.
 FT DOMAIN 22 27 POLY-SER.
 FT CONFICT 259 259 V -> I (IN REF. 1).
 FT CONFLICT 506 506 R -> W (IN REF. 2; CAB38415).
 SQ SEQUENCE 720 AA; 76614 MW; ACD91CCDD00A89C CRC64;
 Query March 7.6%; Score 80; DB 1; Length 720;
 Best Local Similarity 21.1%; Pred. No. 7.5;
 Matches 32; Conservative 21; Mismatches 39; Indels 60; Gaps 6;
 Oy 70 PCRSAPYTELVCSCGQCSPARLLP-----NAIGRGKWRPSPGDFRCIPDRYARV-----121
 Db 341 PANSRPLPSNLSKSGRMGPALRLPALRPGVPGASSW-----QAKRVQVSE 385
 Oy 122 ---QLCPGSEAP-----BARVRLVASC--KCRRLT 148
 Db 386 LAEQLTPAPSPAPQPTPGCGQWINSQAWSESSQLNTRIRRRDSCLNSTKTYMP 445
 Oy 149 RPHNSELKDF-----GTEARPKGRKRPRA 176
 Db 446 TPTNQFKLPKFSIGDSPSTPKLSRAQRFOS 477
 RESULT 14
 SLIT DROME STANDARD; PRT; 1480 AA.
 ID SLIT DROME
 AC P24014;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 15-JUN-2002 (Rel. 21, Last annotation update)
 DE Slit protein precursor.
 GN Slit.
 OS Drosophila melanogaster (fruit fly).
 CC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 CC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 CC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN SEQUENCE FROM N.A.
 RP MEDLINE=9109665; PubMed=2176636;
 RA Rothenberg J.M., Jacobs J.R., Goodman C.S., Artavanis-Tsakonas S.;
 RT "Slit: an extracellular protein necessary for development of midline
 RT glia and commissural axon pathways contains both EGF and LRR
 RT domains.";
 RL Genes Dev. 4:2169-2187(1990).
 CC -1- FUNCTION: NECESSARY FOR DEVELOPMENT OF MIDLINE GLIA AND
 CC COMMISSURAL AXON PATHWAYS. SLIT MAY INTERACT WITH EXTRACELLULAR
 CC MATRIX MOLECULES.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOPFORMS: A LONG FORM (SHOWN HERE) AND A
 CC SHORT FORM, ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: EXCRETED BY THE MIDLINE GLIA CELLS AND
 CC EVENTUALLY DISTRIBUTED ALONG THE AXONS.
 CC -1- SIMILARITY: CONTAINS 7 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 24 LEUCINE-RICH REPEATS (LRR).
 CC -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTEINE KNOT-LIKE (CTCK) DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 LAMININ G-LIKE DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X53959; CAA37910.1; -;
 DR PIR; A36665; A36665.
 DR HSSP; P00740; 1EDM.
 DR FLYBase; FBgn003425; slt.
 DR InterPro; IPR000152; Aex_hydroxyl.
 DR InterPro; IPR000359; Cys_knot.

DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR000742; EGF_2.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR001611; LRR_Nterm.
 DR InterPro; IPR000483; LRR_Cterm.
 DR InterPro; IPR000372; LRR_Nterm.
 DR InterPro; IPR003592; LRR_out.
 DR InterPro; IPR003591; LRR_Typ.
 DR InterPro; IPR001791; Laminin_G.
 DR Pfam; PF00007; Cys_knot; 1.
 DR Pfam; PF00008; EGF_7.
 DR Pfam; PF00054; Laminin_G; 1.
 DR Pfam; PF00560; LRR; 16.
 DR Pfam; PF01462; LRRNT; 4.
 DR Pfam; PF01463; LRRCT; 4.
 DR SMART; SM00041; CT; 1.
 DR SMART; SM00179; EGF_Ca; 2.
 DR SMART; SM00001; EGF_Like; 5.
 DR SMART; SM00370; LRR; 4.
 DR SMART; SM00082; LRRCT; 4.
 DR SMART; SM00013; LRRNT; 4.
 DR SMART; SM00369; LRR_Typ; 9.
 DR SMART; SM00282; LamG; 1.
 DR PROSITE; PS00010; ASX_HYDROXYL; 3.
 DR PROSITE; PS01185; CTCK_1; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 DR PROSITE; PS00022; EGF_1; 7.
 DR PROSITE; PS01186; EGF_2; 5.
 DR PROSITE; PS01187; EGF_Ca; 2.
 DR PROSITE; PS50025; LAM_G_DOMAIN; 1.
 KW Neurogenesis; Glycoprotein; Signal; Alternative splicing;
 KW EGF-like domain; Repeat; Leucine-rich repeat.
 FT CHAIN 1 36
 FT REPEAT 37 1480 SLIT PROTEIN.
 FT REPEAT 99 122 LRR 1.
 FT REPEAT 123 146 LRR 2.
 FT REPEAT 148 170 LRR 3.
 FT REPEAT 171 194 LRR 4.
 FT REPEAT 195 218 LRR 5.
 FT REPEAT 220 246 LRR 6.
 FT REPEAT 321 344 LRR 7.
 FT REPEAT 345 368 LRR 8.
 FT REPEAT 369 392 LRR 9.
 FT REPEAT 394 416 LRR 10.
 FT REPEAT 417 440 LRR 11.
 FT REPEAT 422 445 LRR 12.
 FT REPEAT 527 545 LRR 13.
 FT REPEAT 546 569 LRR 14.
 FT REPEAT 570 593 LRR 15.
 FT REPEAT 595 617 LRR 16.
 FT REPEAT 618 641 LRR 17.
 FT REPEAT 643 666 LRR 18.
 FT REPEAT 678 701 LRR 19.
 FT REPEAT 720 743 LRR 20.
 FT REPEAT 745 764 LRR 21.
 FT REPEAT 765 788 LRR 22.
 FT REPEAT 790 812 LRR 23.
 FT REPEAT 813 836 LRR 24.
 FT REPEAT 838 861 EGF-LIKE 1.
 FT REPEAT 907 944 EGF-LIKE 2.
 FT DOMAIN 946 983 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 985 1022 EGF-LIKE 4.
 FT DOMAIN 1024 1062 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 1064 1100 EGF-LIKE 6.
 FT DOMAIN 1111 1149 LAMININ G-LIKE.
 FT DOMAIN 1152 1325 EGF-LIKE 7.
 FT DOMAIN 1352 1392 EGF-LIKE 7.
 FT DOMAIN 1409 1480 CTCK.
 FT CARBOHYD 111 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 207 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 357 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 435 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 783 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 788 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 958 958 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 998 998 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1060 1060 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1159 1159 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1175 1175 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1243 1243 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1292 1292 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT DISULFID 911 922 BY SIMILARITY.
 FT DISULFID 916 932 BY SIMILARITY.
 FT DISULFID 934 943 BY SIMILARITY.
 FT DISULFID 950 961 BY SIMILARITY.
 FT DISULFID 955 971 BY SIMILARITY.
 FT DISULFID 973 982 BY SIMILARITY.
 FT DISULFID 989 1001 BY SIMILARITY.
 FT DISULFID 995 1010 BY SIMILARITY.
 FT DISULFID 1012 1021 BY SIMILARITY.
 FT DISULFID 1028 1041 BY SIMILARITY.
 FT DISULFID 1035 1050 BY SIMILARITY.
 FT DISULFID 1052 1061 BY SIMILARITY.
 FT DISULFID 1068 1079 BY SIMILARITY.
 FT DISULFID 1073 1088 BY SIMILARITY.
 FT DISULFID 1090 1099 BY SIMILARITY.
 FT DISULFID 1115 1125 BY SIMILARITY.
 FT DISULFID 1120 1137 BY SIMILARITY.
 FT DISULFID 1139 1148 BY SIMILARITY.
 FT DISULFID 1357 1368 BY SIMILARITY.
 FT DISULFID 1362 1380 BY SIMILARITY.
 FT DISULFID 1382 1391 BY SIMILARITY.
 FT DISULFID 1409 1443 BY SIMILARITY.
 FT DISULFID 1423 1457 BY SIMILARITY.
 FT DISULFID 1434 1473 BY SIMILARITY.
 FT DISULFID 1438 1475 BY SIMILARITY.
 FT DISULFID 1442 1479 BY SIMILARITY.
 FT VARSPLIC 1394 1404 MISSING (IN SHORT ISOFORM).
 SQ SEQUENCE 1480 AA; 165752 MW; F9D5925FC170B1C3 CRC64;

Query Match 7.6%; Score 80; DB 1; Length 1480;
 Best Local Similarity 18.0%; Pred. No. 16;
 Matches 31; Conservative 23; Mismatches 62; Indels 56; Gaps 5;

QY 4 QAFKDATEIIPLEGVEYPP--PELEN-----NKTNN 34
 DB 1333 QEEDEDEDQFMDETPHIKEEVDPLENCKRGSRGVNSNARDGYQCKKHGGRGYCD 1392
 QY 35 RAENGRPPHPFETKDVSEYSCRELHFTRYVTDPGCSAKVTELVCSGGCGPARLLPN 94
 DB 1393 QGEGSTPP-----TVTAATCKKEQVREYTYTENDCRSRQPLTKACVGGCG----- 1439
 QY 95 AIGRGKWRPSPGPPFCIPDRYRAQRYGLCPGGEAPPARKRLVASCKCKR 146
 DB 1440 -----NCCCAKIVRRKRYMVC-SNNRKYIKLIVRKCGCTK 1477

RESULT 15
 NK2E_MOUSE STANDARD; PRT; 318 AA.
 AC P42582; P97335;
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 41, Last annotation update)
 DT 13-UN-2002 (Rel. 1, Last sequence update)
 DE Homeobox protein NK-2 homolog E (Homeobox protein NKX-2.5) (Cardiac-specific homeobox) (Homeobox protein CSX).
 GN NKX2E OR CSX OR NKX2-5 OR NKX-2.5.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Tissue=Heart;
 RX MEDLINE=94116443; PubMed=7904557;
 RA Lints T.J., Parsons L.M., Hartley L., Lyons I., Harvey R.P.;
 RT "NKX-2.5: a novel murine homeobox gene expressed in early heart

1
2
3
4

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:12:16 ; Search time 54.72 Seconds
(without alignments)
715.441 Million cell updates/sec

Title: US-09-867-274-2

Perfect score: 1049
Sequence: 1 QGMQAFKNDATETIPELGEY.....KRPFRASAKANQALENAY 190

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_virus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match length	DB	ID	Description
1	344	32.8	206	11	Q9CQN4	Q9CQN4 mus musculus
2	341	32.5	206	4	Q96HJ7	Q96HJ7 homo sapien
3	270	25.7	134	4	Q9Y3U3	Q9Y3U3 homo sapien
4	106	10.1	272	13	Q9PW80	Q9PW80 gallus galli
5	104	9.9	272	13	Q9PUK2	Q9PUK2 gallus galli
6	90.5	8.6	168	4	Q9H772	Q9H772 homo sapien
7	89.5	8.5	184	4	Q60565	Q60565 homo sapien
8	89.5	8.5	184	6	Q8WNY1	Q8WNY1 macaca mulia
9	89.5	8.5	215	8	Q9NFB6	Q9NFB6 beta vulgar
10	88.5	8.4	168	11	Q88273	Q88273 mus musculus
11	88	8.4	1114	11	Q9KXW7	Q9KXW7 mus musculus
12	87.5	8.3	184	13	Q73755	Q73755 gallus galli
13	87.5	8.3	558	4	Q96D88	Q96D88 homo sapien
14	87.5	8.3	585	4	Q9BNM1	Q9BNM1 homo sapien
15	87.5	8.3	585	4	Q8WXX2	Q8WXX2 homo sapien
16	86.5	8.2	182	13	Q73754	Q73754 xenopus lae

17	86	8.2	184	11	Q35793	Q35793 rattus norv
18	85	8.2	184	11	Q70326	Q70326 mus musculus
19	85	8.1	181	4	Q96168	Q96168 homo sapien
20	83	7.9	958	5	Q9V560	Q9V560 drosophila
21	82.5	7.9	322	4	Q96FP7	Q96FP7 homo sapien
22	82.5	7.9	329	4	Q9W60	Q9W60 acetobacter
23	82	7.8	141	6	Q95J85	Q95J85 monodelphis
24	82	7.8	988	6	Q97867	Q97867 sus scrofa
25	81.5	7.8	661	5	Q9V7U8	Q9V7U8 drosophila
26	81.5	7.8	669	5	Q8T0D3	Q8T0D3 drosophila
27	81	7.7	270	13	P70041	P70041 xenopus lae
28	81	7.7	1096	10	Q9X524	Q9X524 oryza sativ
29	80.5	7.7	267	4	Q95813	Q95813 homo sapien
30	80.5	7.7	715	4	Q9Y4Q3	Q9Y4Q3 homo sapien
31	80	7.6	219	12	Q91T11	Q91T11 tupia leup
32	80	7.6	258	4	Q8FAT2	Q8FAT2 homo sapien
33	80	7.6	712	4	Q9V557	Q9V557 homo sapien
34	80	7.6	720	4	Q9N123	Q9N123 homo sapien
35	80	7.6	720	4	Q9BRE0	Q9BRE0 homo sapien
36	80	7.6	891	10	Q8RYT6	Q8RYT6 oryza sativ
37	80	7.6	1440	5	Q20204	Q20204 caenorhabdi
38	80	7.6	1480	5	Q9V7P8	Q9V7P8 drosophila
39	80	7.6	1504	5	Q9XYV4	Q9XYV4 drosophila
40	80	7.6	1504	5	Q9V7P9	Q9V7P9 drosophila
41	79.5	7.6	318	11	Q925V3	Q925V3 mus musculus
42	79.5	7.6	386	10	Q8S5P2	Q8S5P2 oryza sativ
43	79.5	7.6	500	10	Q9M6C2	Q9M6C2 trifolium r
44	79.5	7.6	1081	4	Q76065	Q76065 homo sapien
45	79	7.5	220	10	Q9PP95	Q9PP95 oryza sativ

ALIGNMENTS

RESULT 1
ID Q9CQN4 PRELIMINARY; PRT; 206 AA.
AC Q9CQN4;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE 0610006G05RIK protein (RIKEN CDNA 0610006G05 gene).
GN 0610006G05RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PANCREAS; AND KIDNEY;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamada I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochaya H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Straubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barns G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Guelinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohatsu S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 403:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.

RL Curr. Biol. 9:931-938(1999).
 DR EMBL; AF139721; AAD51610.1; -
 DR InterPro; IPR000359; Cys_knot.
 DR InterPro; IPR004133; DAN_dom.
 DR Pfam; PF03045; DAN; 1.
 DR SMART; SM00041; CT; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 SQ SEQUENCE 272 AA; 31224 MW; 36E4C9F719711BCA CRC64;

Query Match 10.1%; Score 106; DB 13; Length 272;
 Best Local Similarity 23.7%; Pred. No. 0.0066;
 Matches 32; Conservative 20; Mismatches 57; Indels 26; Gaps 5;

QY 22 EPPPELENNK---TWNRALNGRPPHPHFTKDVSEVSCRELAHFRVYTDGPRSAKPV 77
 DB 129 EPPYRKDAKKFMDHFLRKNSASBEVVLPIKTENMHQETCTLPFSQSVAHSCCKV-IV 187
 QY 78 TELVSGGCGPARLLPNAIGRGKMRPSPDFR-----CIPDRYARQVQLCPGGEA 130
 DB 188 QNNLCFGKCSS-----FHVPGPDDRLYTFCKSKCLPTKFSMGLDNLCT-SSV 233
 QY 131 PRARKRLVASCKK 145
 DB 234 PVKAKMIVECNCE 248

RESULT 5

Q9PUK2 PRELIMINARY; PRT; 272 AA.

AC Q9PUK2;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Caronte.
 GN CAR.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Rodriguez Esteban C., Capdevila J., Economides A.N., Pascual J.,
 RA Ortiz A., Izpisua Belmonte J.C.;
 RT "Caronte, a novel cer-like protein, mediates the establishment of
 RT embryonic left-right asymmetry.";
 RL Nature 0:0-0(1999).
 DR EMBL; AF179484; AAD55581.1; -
 DR InterPro; IPR000359; Cys_knot.
 DR InterPro; IPR004133; DAN_dom.
 DR Pfam; PF03045; DAN; 1.
 DR SMART; SM00041; CT; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 SQ SEQUENCE 272 AA; 31201 MW; 0D89729715771BC2 CRC64;

Query Match 9.9%; Score 104; DB 13; Length 272;
 Best Local Similarity 23.7%; Pred. No. 0.011;
 Matches 32; Conservative 19; Mismatches 58; Indels 26; Gaps 5;

QY 22 EPPPELENNK---TWNRALNGRPPHPHFTKDVSEVSCRELAHFRVYTDGPRSAKPV 77
 DB 129 EPPYRKDAKKFMDHFLRKNSASBEVVLPIKTENMHQETCTLPFSQSVAHSCCKV-IV 187
 QY 78 TELVSGGCGPARLLPNAIGRGKMRPSPDFR-----CIPDRYARQVQLCPGGEA 130
 DB 188 QNNLCFGKCSS-----FHVPGPDDRLYTFCKSKCLPTKFSMGLDNLCT-SSV 233
 QY 131 PRARKRLVASCKK 145
 DB 234 PVKAKMIVECNCE 248

RESULT 6

Q9H772 PRELIMINARY; PRT; 168 AA.

AC Q9H772;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE CDNA: FLJ21195 f18, clone COL00185.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=COLON;
 RA Kawabata A., Hiki T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
 RA Tanaka T., Nakamura Y., Iisgai T., Sugano S.;
 RT "NEDO human cDNA sequencing project.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK024848; BAB15026.1; -
 DR InterPro; IPR000359; Cys_knot.
 DR InterPro; IPR004133; DAN_dom.
 DR Pfam; PF03045; DAN; 1.
 DR SMART; SM00041; CT; 1.
 DR PROSITE; PS01225; CTCK_2; 1.
 SQ SEQUENCE 168 AA; 19320 MW; D5A4E818BF8C0E CRC64;

Query Match 8.6%; Score 90.5; DB 4; Length 168;
 Best Local Similarity 26.5%; Pred. No. 0.16;
 Matches 27; Conservative 19; Mismatches 43; Indels 13; Gaps 5;

QY 48 ETKVSEVSCRELAHFRVYTDGPRSAKPVTELVCSCGCGPARLLPNAIGRGKMRPSPG 107
 DB 64 ERKYKSDWCKTQPLRGTVSESGRS--RTILNRFCYQGN-SFIPRHV-----KKEE 115
 QY 108 DFR---CIPDRYARQVQLCPGGEAP-PRARKRLVASCKK 144
 DB 116 SFOSCAFCKPQRTVSVVLEECPLDPPFLKTKIOVKVGCRC 157

RESULT 7

ID 060565 PRELIMINARY; PRT; 184 AA.

AC 060565;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE GREMLIN (DRM).
 GN DRM.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98325381; PubMed=960951;
 RA Hau D.R., Economides A.N., Wang X., Eimon P.M., Harland R.M.;
 RT "The Xenopus dorsalizing factor Gremlin identifies a novel family of
 RT secreted proteins that antagonize BMP activities.";
 RL Mol. Cell 1:673-683(1998).
 RN (2)
 RP SEQUENCE FROM N.A.
 RA Tate G., Mitsuya T.;
 RT "Human Gremlin homologue";
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RN (3)
 RP SEQUENCE FROM N.A.
 RA McMahon R.A., Murphy M., Clarkson M.R., Godson C., Martin F.,
 RA Brady H.R.;
 RT "HIG-2, a mesangial cell gene differentially induced in high glucose,
 RT is human gremlin.";
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 RN (4)

RP SEQUENCE FROM N.A.
 RC TISSUE=SMALL INTESTINE;
 RX MEDLINE=2035496; PubMed=10894942;
 RA Topol L.Z., Modi W.S., Koochekpour S., Blair D.G.;
 RT "DRM/GREVLIN (CTSF1B1) maps to human chromosome 15 and is highly
 expressed in adult and fetal brain.";
 RL Cytogenet. Cell Genet. 89:79-84(2000).
 DR EMBL: AF045800; AAC39725.1; -
 DR EMBL: AB032372; BA84462.1; -
 DR EMBL: AF110137; AA06677.1; -
 DR EMBL: AF154054; AAG23891.1; -
 DR InterPro: IPR000359; Cys_knot.
 DR InterPro: IPR004133; DAN_dom.
 DR Pfam: PF03045; DAN_1.
 DR SMART: SM00041; CT; 1.
 SQ SEQUENCE 184 AA; 20697 MW; 4B588598DE12C47E CRC64;

Query Match 8.5%; Score 89.5; DB 4; Length 184;
 Best Local Similarity 24.1%; Pred. No. 0.22;
 Matches 38; Conservative 22; Mismatches 57; Indels 41; Gaps 9;

QY 18 GEYEPPEPELENNK-----TNRRAENGRRPHHPETKDYSEYSCRELHFT--R 64
 DB 31 GAIP-PPDKAQNDSQTPQDPGRGGRGTAMGE--EVLSSQELHHTERK 87
 QY 65 YVTDGPR-----SAKPVELVCSGCGCPARLLPNAIGRKWRRSPGDFR-110
 DB 88 YLRKDWCKTQPLKQTIHEBGNSTIINRCYGCN-SFIIPRI-----RKEGSGFS 140

QY 111 --CIPDRYRAQRYVOLLCPGGEAP-RARKYRLVASCKC 144
 DB 141 CSFCKPKKFTTMMVTLNCPELQPTKKRVTRVQKRC 178

RESULT 8
 ID Q8MNY1 PRELIMINARY; PRT; 184 AA.
 AC Q8MNY1;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Gremilin.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Macaca.
 OC NCBI_TaxID=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Christenson L.K., Duffey D.M.;
 RL Submitted (JCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF39783; ALJ32022.1; -
 DR InterPro: IPR000359; Cys_knot.
 DR InterPro: IPR004133; DAN_dom.
 DR Pfam: PF03045; DAN_1.
 DR SMART: SM00041; CT; 1.
 SQ SEQUENCE 184 AA; 20697 MW; 4B588598DE12C47E CRC64;

Query Match 8.5%; Score 89.5; DB 6; Length 184;
 Best Local Similarity 24.1%; Pred. No. 0.22;
 Matches 38; Conservative 22; Mismatches 57; Indels 41; Gaps 9;

QY 18 GEYEPPEPELENNK-----TNRRAENGRRPHHPETKDYSEYSCRELHFT--R 64
 DB 31 GAIP-PPDKAQNDSQTPQDPGRGGRGTAMGE--EVLSSQELHHTERK 87
 QY 65 YVTDGPR-----SAKPVELVCSGCGCPARLLPNAIGRKWRRSPGDFR-110
 DB 88 YLRKDWCKTQPLKQTIHEBGNSTIINRCYGCN-SFIIPRI-----RKEGSGFS 140
 QY 111 --CIPDRYRAQRYVOLLCPGGEAP-RARKYRLVASCKC 144
 DB 141 CSFCKPKKFTTMMVTLNCPELQPTKKRVTRVQKRC 178

RESULT 9
 ID Q9MFB6 PRELIMINARY; PRT; 215 AA.
 AC Q9MFB6;
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE Orf215 protein.
 GN ORF215.
 OS Beta vulgaris (Sugar beet).
 OS Mitochondrion.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Beta.
 OC NCBI_TaxID=3555;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=TK81-O.
 RX MEDLINE=20330382; PubMed=10871408;
 RA Kubo T., Nishizawa S., Sugawara A., Itchoda N., Estiati A., Mikami T.;
 RT "The complete nucleotide sequence of the mitochondrial genome of sugar
 beet (Beta vulgaris L.) reveals a novel gene for tRNA^{Cys} (GCA).";
 RL Nucleic Acids Res. 28:2571-2576(2000).
 DR EMBL: AP000396; BAA9326.1; -
 KW Mitochondrion.
 SQ SEQUENCE 215 AA; 24100 MW; C144AE9AB41E5E06 CRC64;

Query Match 8.5%; Score 89.5; DB 8; Length 215;
 Best Local Similarity 28.7%; Pred. No. 0.26;
 Matches 31; Conservative 14; Mismatches 32; Indels 31; Gaps 9;

QY 3 WQAFKQDATEITIPELGEPPEPELENNKTKMRAENGRRPHHPETKDYSEYSCRELHF 62
 DB 134 WQGLINDS-----PEGDAVSDAENVPAEN--PGHNEAPAPAP-NCR51H-178
 QY 63 TRYTVDGPRSAKPYTELVCSCGCPARLLPN--AIGRKW-WRSPG 107
 DB 179 -----PCNG--PWTRPIC-GNC---FLIINWIKITIKGRWNRKGP 213

RESULT 10
 ID O88273 PRELIMINARY; PRT; 168 AA.
 AC O88273;
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE PRDC.
 GN PRDC OR PRDC.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=C57BL6; TISSUE=CNS;
 RX MEDLINE=96301233; PubMed=9639362;
 RA Minabe-Saegusa C., Saegusa H., Tsukahara M., Noguchi S.;
 RT "Sequence and expression of a novel mouse gene PRDC (protein related
 to DAN and cerberus) identified by a gene trap approach.";
 RT Dev. Growth Differ. 40:343-353(1998).
 DR EMBL: AB011030; BAA29038.1; -
 DR MGD; WGI:1344367; Prdc.
 DR InterPro: IPR000359; Cys_knot.
 DR InterPro: IPR004133; DAN_dom.
 DR Pfam: PF03045; DAN_1.
 DR SMART: SM00041; CT; 1.
 DR PROSITE: PS01225; CTCK_2; 1.
 SQ SEQUENCE 168 AA; 19334 MW; 6361C1581D49C281 CRC64;

Query Match 8.4%; Score 88.5; DB 11; Length 168;

Best Local Similarity 26.5%; Pred. No. 0.25;
Matches 27; Conservative 19; Mismatches 43; Indels 13; Gaps 5;

QY 48 ETDVSEYSGREHFRVYTDGCRSAKPTVELVCGGCGPALLPNAIGRGKMWPSGP 107
DB 64 ERYKLSDWCKTLPLOTJSEGCSS-RITLNFRCYGCQV-SFYIPRHV-----KKEED 115
QY 108 DFR-----CIPDRYAROVOLLCGSGEAP-PARKYRVASCKC 144
DB 116 SFQSCAFCKRQRTSVIVLEECGDPDPFRKIKIKYKHCRC 157

RESULT 11

09JWK7

PRELIMINARY; PRT; 1114 AA.

AC 09JWK7; PRELIMINARY; PRT; 1114 AA.
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Cysteine-rich protein NFX-1.
GN NFX1 OR 1300017N15RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

RA Arlootta P., Jenkins N.A., Jay G., Ono S.J.;
RT "Isolation of a Full-length Murine NFX-1 cDNA, its Chromosomal
Location and Developmental Expression."
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF223576; AAF34700.1; -
DR MGI; 1921414; NFX1.
DR InterPro; IPR001374; R3H.
DR InterPro; IPR000967; Znf_NFX1.
DR InterPro; IPR001965; Znf_PHD.
DR InterPro; IPR001841; Znf_Ring.
DR Pfam; PF01424; R3H; 1.
DR Pfam; PF01422; ZF-NFX1; 8.
DR SMART; SM00393; R3H; 1.
DR SMART; SM00184; RING; 1.
DR SMART; SM00438; ZNF_NFX; 9.
SQ SEQUENCE 1114 AA; 123878 MW; 857C90260D28EA4B CRC64;

Query Match 8.4%; Score 88; DB 11; Length 1114;
Best Local Similarity 24.5%; Pred. No. 2.2; Indels 34; Gaps 9;

Matches 46; Conservative 22; Mismatches 86; Indels 34; Gaps 9;

QY 1 QGNQAFKND-----ATEIIPELGEYPPPELENNKTMNRAENGGR-PHPPEFKDVSEY 55
DB 101 QPNQKARNEHQNRKKAQGLSEGTSTSLF---SVARSSEGTNPREHSPSE----- 152
QY 56 SCREHETRYVTDGCRSAKPTVELVCGGCGPALLPNAIG--RGKMWPSGDPDFRC-1 112
DB 153 -----KEVVDLDPGAKPRKKAQLTINYGRG--PKAKGRSRSEGRMSPKSEDEI 201
QY 113 PDYRAQROVOLLCGSGEAPARKYRVASCKCRITRFHNSQLKDGCTAARPOKGRKP 172
DB 202 PDYRAQROVOLLCGSGEAPARKYRVASCKCRITRFHNSQLKDGCTAARPOKGRKP 255
QY 173 RPRARSASAK 180
DB 256 -PKQESQR 262

RESULT 12

073755

PRELIMINARY; PRT; 184 AA.

AC 073755; PRELIMINARY; PRT; 184 AA.
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Gremilin.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archaeosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.

OX NCBI_TaxID=9031;

RN [1]
RP SEQUENCE FROM N.A.

RX MEDLINE=98325381; PubMed=9660951;

RA Hsu D.R., Economides A.N., Wang X., Elmon P.M., Harland R.M.;

RT "The Xenopus dorsalizing factor Gremilin identifies a new family of

secreted proteins that antagonize BMP activities."

RL Mol. Cell 1:673-683(1998).

DR EMBL; AF045799; AAC41280.1; -

DR InterPro; IPR000359; Cys_knot.

DR InterPro; IPR004133; DAN_dom.

DR Pfam; PF03045; DAN; 1.

DR SMART; SM00041; CT; 1.

SQ SEQUENCE 184 AA; 21166 MW; 3510B44B8F6DD5EA CRC64;

Query Match 8.3%; Score 87.5; DB 13; Length 184;
Best Local Similarity 22.4%; Pred. No. 0.35;

Matches 34; Conservative 24; Mismatches 55; Indels 39; Gaps 8;

QY 23 PPELE-----NNKTMNRAENGGRPPHPPEFKDVSEYSGREHFT--RYVTDGP 70
DB 36 PPPDKQDPDSEOMQTOOQSGSRHREGRGTSMPAE--EVLSSQEGALHITERKYLKRDW 93
QY 71 CR-----SAPVTELVCGGCGPALLPNAIGRGKMWPSGDPDFR-----CIP 113
DB 94 CTQPLKQTHIEGGSNRTIINFRCYGCQV-SFYIPRHV-----RKEGSGFCSGFCRP 146
QY 114 DRYRAQROVOLLCGSGEAPARKYRVASCKC 144
DB 147 KFTMTVTVLNCPQLPQPRKKRITRVKRCRC 178

RESULT 13

096D88

PRELIMINARY; PRT; 558 AA.

AC 096D88; PRELIMINARY; PRT; 558 AA.
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Similar to hypothetical protein MG3121 (Fragment).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

OX NCBI_TaxID=9606;

RN [1]
RP SEQUENCE FROM N.A.

RC TISSUE=CERVIX;

RA Strauberg R.;

RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC010232; AAH10232.1; -

KW Hypothetical protein.

FT NON TER 1

SQ SEQUENCE 558 AA; 60917 MW; FD48933F9F3DB5B6 CRC64;

Query Match 8.3%; Score 87.5; DB 4; Length 558;
Best Local Similarity 23.2%; Pred. No. 1.2;

Matches 46; Conservative 14; Mismatches 45; Indels 93; Gaps 10;

QY 21 PPEPELENNKTMNRAENG-----RPPHPPEFKDVS-----EYSCR 58
DB 249 PSPPMKLEIKIAISAEQGAEGTASVSPRPIROWTODINTTALLPKPSLGRSYSCP 308
QY 59 ELHFTRYVTDGCRSAKPTVELVCGGCGPALLPNAIGRGKMWPSGDPDFRCIPDRYRA 118
DB 309 DL-----GPPPGTCT-----WPPAPDQ-----PSRPP 332
QY 119 QRVOLLCGGEAPRA-----RK-----VRLVASCCKRLTRFHNSQLKDF 159
DB 333 RRTYV--GGEMARAPPPPPCLRKVEVPLGGVGASPSLTISCSSTASTSP----- 381

QY 160 GTEARPO---KGRKR 173
D 382 -SEPAEPRLGSTKGKPR 398

RESULT 14

Q9BWN1 PRELIMINARY; PRT; 585 AA.
AC Q9BWN1;
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Hypothetical 64.3 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=EYE;
RA Strausberg R.;
RL Submitted (NCV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC000119; AAH0119.1; -
DR InterPro: IPR002965; P-rich_extensn.
DR PRINTS; PRO1217; PRICHEXTENSN.
KW Hypothetical protein.
SQ SEQUENCE 585 AA; 64328 MW; EE706853CE2D2742 CRC64;

Query March 8.3%; Score 87.5; DB 4; Length 585;
Best Local Similarity 23.2%; Pred. No. 1.2;
Matches 46; Conservative 14; Mismatches 45; Indels 93; Gaps 10;

QY 21 PEPPPELENNKTNRAENG-----RPPHPFETKDV-----EYSCR 58
D 276 PSEPMKLELKIAISEAEAGSAGASVSPRPPIQWRTQDHTPALPKSLGRSISCP 335
QY 59 ELHFTRYVTDGPCRSAPVTELVCSGQCGPARLLPNAIGRGKMWPPSGDPDFRCIPDRYRA 118
D 336 DL-----GPPGPGTCT-----WPPAPPO-----PSRRL 359
QY 119 QRVQLLCPGGEAPRA-----RK-----VRLVASCCKRLTRFNOSELKDF 159
D 360 RRTTV--GGGEWAPPPPPRCLKEVPLGAGVGSPLTTSSTASTSF----- 408
QY 160 GTEARPO---KGRKR 173
D 409 -SEPAEPRLGSTKGKPR 425

RESULT 15

Q8WTK2 PRELIMINARY; PRT; 585 AA.
AC Q8WTK2;
DT 01-MAR-2002 (TREMblrel. 20, Created)
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Hypothetical 64.3 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=RETINA;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC021934; AAH21934.1; -
KW Hypothetical protein.
SQ SEQUENCE 585 AA; 64344 MW; 2DF54B1551CE4C CRC64;

Query Match 8.3%; Score 87.5; DB 4; Length 585;
Best Local Similarity 23.2%; Pred. No. 1.2;

Matches 46; Conservative 14; Mismatches 45; Indels 93; Gaps 10;
QY 21 PEPPPELENNKTNRAENG-----RPPHPFETKDV-----EYSCR 58
D 276 PSEPMKLELKIAISEAEAGSAGASVSPRPPIQWRTQDHTPALPKSLGRSISCP 335
QY 59 ELHFTRYVTDGPCRSAPVTELVCSGQCGPARLLPNAIGRGKMWPPSGDPDFRCIPDRYRA 118
D 336 DL-----GPPGPGTCT-----WPPAPPO-----PSRRL 359
QY 119 QRVQLLCPGGEAPRA-----RK-----VRLVASCCKRLTRFNOSELKDF 159
D 360 RRTTV--GGGEWAPPPPPRCLKEVPLGAGVGSPLTTSSTASTSF----- 408
QY 160 GTEARPO---KGRKR 173
D 409 -SEPAEPRLGSTKGKPR 425

Search completed: March 28, 2003, 14:18:46
Job time : 57.72 secs

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:16:22 ; Search time 12.16 Seconds
(without alignments)
917.557 Million cell updates/sec

Title: US-09-867-274-2
Perfect score: 1049
Sequence: 1 QGMQAFKNDATRIIPELGEY.....KPRPARSAKANOAELENNY 190

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 237916 seqs, 58723674 residues
Total number of hits satisfying chosen parameters: 237916

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US05_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1049	100.0	190	10	US-09-867-274-2
2	1049	100.0	213	10	US-09-867-274-5
3	932.5	88.9	185	10	US-09-867-274-4
4	932.5	88.9	208	10	US-09-867-274-6
5	758	72.3	139	9	US-09-864-761-47109
6	344	32.8	206	9	US-09-866-050A-159
7	344	32.8	206	9	US-09-866-050A-286
8	341	32.5	206	10	US-09-853-625B-2
9	341	32.5	206	10	US-09-853-625B-20
10	337	32.1	183	10	US-09-867-274-25
11	98	9.3	337	10	US-09-814-777A-126
12	89.5	8.5	184	9	US-10-159-749-2
13	89.5	8.5	184	9	US-10-159-749-6
14	87.5	8.3	184	9	US-10-159-749-4
15	86.5	8.2	182	9	US-10-159-749-9
16	86	8.2	184	9	US-10-159-749-9
17	85	8.1	181	9	US-10-159-749-8
18	82.5	7.9	462	10	US-09-814-777A-99
19	82	7.8	150	10	US-09-925-301-1324

20	81	7.7	270	9	US-09-903-170C-1	Sequence 1, Appli
21	81	7.7	270	10	US-09-903-180B-1	Sequence 1, Appli
22	81	7.7	270	10	US-09-903-187A-1	Sequence 1, Appli
23	81	7.7	270	10	US-09-903-471A-1	Sequence 1, Appli
24	81	7.7	270	10	US-09-903-188A-1	Sequence 1, Appli
25	81	7.7	270	10	US-09-903-323A-1	Sequence 1, Appli
26	81	7.7	270	10	US-09-903-325A-1	Sequence 1, Appli
27	80.5	7.7	267	9	US-10-044-716-12	Sequence 12, Appli
28	80.5	7.7	267	9	US-09-089-818B-8	Sequence 8, Appli
29	80	7.6	1480	12	US-10-011-064-5	Sequence 5, Appli
30	79	7.5	336	10	US-09-756-186-8	Sequence 8, Appli
31	78	7.4	1210	10	US-09-725-433-2	Sequence 2, Appli
32	77	7.3	510	10	US-09-866-562-56	Sequence 56, Appli
33	76.5	7.3	372	9	US-10-029-180-10	Sequence 10, Appli
34	76	7.2	601	10	US-09-925-301-844	Sequence 844, App
35	75.5	7.2	1436	9	US-10-042-431-78	Sequence 78, Appli
36	75.5	7.2	1436	9	US-09-759-130B-448	Sequence 48, App
37	75	7.1	4123	9	US-10-213-509-5	Sequence 5, Appli
38	73.5	7.0	400	9	US-10-026-021-5	Sequence 249, App
39	73.5	7.0	685	10	US-09-771-161A-249	Sequence 250, App
40	73.5	7.0	685	10	US-09-771-161A-250	Sequence 251, App
41	73.5	7.0	685	10	US-09-771-161A-251	Sequence 250, App
42	72.5	6.9	104	10	US-09-764-887-260	Sequence 260, App
43	72.5	6.9	247	9	US-10-050-704-275	Sequence 275, Appli
44	72.5	6.9	547	9	US-09-977-418-4	Sequence 4, Appli
45	72	6.9	251	10	US-09-764-853-522	Sequence 522, App

ALIGNMENTS

RESULT 1
US-09-867-274-2
Sequence 2, Application US/09867274
Patent No. US2002010650A1
GENERAL INFORMATION:
APPLICANT: Paszty, Christopher
APPLICANT: Gao, Yongming
TITLE OF INVENTION: Cysteine knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
FILE REFERENCE: 01017/37428
CURRENT APPLICATION NUMBER: US/09/867, 274
CURRENT FILING DATE: 2001-05-29
PRIOR APPLICATION NUMBER: US 60/208, 550
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: US 60/223, 542
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 190
TYPE: PRT
ORGANISM: Homo sapiens
US-09-867-274-2

Query Match 100.0%; Score 1049; DB 10; Length 190;
Best Local Similarity 100.0%; Pred. No. 4.8e-96;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	QGMQAFKNDATRIIPELGEYPPPELENNKTNFAENGRRPHHPFTKDVSEYSCREL	60
DB	1	QGMQAFKNDATRIIPELGEYPPPELENNKTNFAENGRRPHHPFTKDVSEYSCREL	60
QY	61	HTTRVYTGDPCHSAPVETLVSGGCGPARLLPNAIGKMKWRPSPGPPRCIPDYYRQR	120
DB	61	HTTRVYTGDPCHSAPVETLVSGGCGPARLLPNAIGKMKWRPSPGPPRCIPDYYRQR	120
QY	121	VOLLCPGSGAPPARKRVRLVASCCKRLTRFNQSEIKDFTGEAARPOKGRKPRPARSAK	180
DB	121	VOLLCPGSGAPPARKRVRLVASCCKRLTRFNQSEIKDFTGEAARPOKGRKPRPARSAK	180
QY	181	ANOAELENNY 190	
DB	181	ANOAELENNY 190	

RESULT 2
US-09-867-274-5
; Sequence 5, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; APPLICANT: Gao, Yongming
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-867-274-5
Query Match 100.0%; Score 1049; DB 10; Length 213;
Best Local Similarity 100.0%; Pred. No. 5.5e-96;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 OGMQAFKNDATETIIPELGEYEPPEPELENNKTNRAENGRRPHHPETKDVSEYSCREL 60
DB 24 OGMQAFKNDATETIIPELGEYEPPEPELENNKTNRAENGRRPHHPETKDVSEYSCREL 83
QY 61 HFTRYVTDGCRSAKPVTETVCSGCCGPARRLLPVAIGRGKWMRPSGDPFCIPDRYRAOR 120
DB 84 HFTRYVTDGCRSAKPVTETVCSGCCGPARRLLPVAIGRGKWMRPSGDPFCIPDRYRAOR 143
QY 121 VOLCPGGEAPRAKRVLVASCKCKRLTRFNOSSELDGFEARPOKGRKPRPARSAK 180
DB 144 VOLCPGGEAPRAKRVLVASCKCKRLTRFNOSSELDGFEARPOKGRKPRPARSAK 203
QY 181 ANQAELENAY 190
DB 204 ANQAELENAY 213
RESULT 3
US-09-867-274-4
; Sequence 4, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; APPLICANT: Gao, Yongming
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-867-274-4
Query Match 88.9%; Score 932.5; DB 10; Length 185;
Best Local Similarity 88.4%; Pred. No. 1.4e-84;
Matches 168; Conservative 10; Mismatches 7; Indels 5; Gaps 2;

QY 1 OGMQAFKNDATETIIPELGEYEPPEPELENNKTNRAENGRRPHHPETKDVSEYSCREL 60
DB 1 OGMQAFKNDATETIIPELGEYEPPEPELENNKTNRAENGRRPHHPETKDVSEYSCREL 58
QY 61 HFTRYVTDGCRSAKPVTETVCSGCCGPARRLLPVAIGRGKWMRPSGDPFCIPDRYRAOR 120
DB 59 HFTRYVTDGCRSAKPVTETVCSGCCGPARRLLPVAIGRGKWMRPSGDPFCIPDRYRAOR 118
QY 121 VOLCPGGEAPRAKRVLVASCKCKRLTRFNOSSELDGFEARPOKGRKPRPARSAK 180
DB 119 VOLCPGGEAPRAKRVLVASCKCKRLTRFNOSSELDGFEARPOKGRKPRPARSAK 175
QY 181 ANQAELENAY 190
DB 176 ANQAELENAY 185
RESULT 4
US-09-867-274-6
; Sequence 6, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; APPLICANT: Gao, Yongming
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-867-274-6
Query Match 88.9%; Score 932.5; DB 10; Length 208;
Best Local Similarity 88.4%; Pred. No. 1.7e-84;
Matches 168; Conservative 10; Mismatches 7; Indels 5; Gaps 2;
QY 1 OGMQAFKNDATETIIPELGEYEPPEPELENNKTNRAENGRRPHHPETKDVSEYSCREL 60
DB 24 OGMQAFKNDATETIIPELGEYEPPEPELENNKTNRAENGRRPHHPETKDVSEYSCREL 81
QY 61 HFTRYVTDGCRSAKPVTETVCSGCCGPARRLLPVAIGRGKWMRPSGDPFCIPDRYRAOR 120
DB 82 HFTRYVTDGCRSAKPVTETVCSGCCGPARRLLPVAIGRGKWMRPSGDPFCIPDRYRAOR 141
QY 121 VOLCPGGEAPRAKRVLVASCKCKRLTRFNOSSELDGFEARPOKGRKPRPARSAK 180
DB 142 VOLCPGGEAPRAKRVLVASCKCKRLTRFNOSSELDGFEARPOKGRKPRPARSAK 198
QY 181 ANQAELENAY 190
DB 199 ANQAELENAY 208
RESULT 5
US-09-864-761-47109
; Sequence 47109, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1

```

; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47109
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC003098.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.62
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.49
; OTHER INFORMATION: SWISSPROT HIT: P45646, EVALU 4.70e-01
US-09-864-761-47109
```

```

Query Match 72.3%; Score 758; DB 10; Length 139;
Best Local Similarity 100.0%; Pred. No. 1.7e-67;
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 52 VSEYSGRELFHTYVTDGPRSAKPYTELVCSCGCGPARLLPNAIGRGKWRSPGDFRC 111
DB 1 VSEYSGRELFHTYVTDGPRSAKPYTELVCSCGCGPARLLPNAIGRGKWRSPGDFRC 60

QY 112 IPDPRYQAVQVLLCPGGEAPRAKRVLVASCKKRLTRFNOSLKDFTGEARPOKGRK 171
DB 61 IPDPRYQAVQVLLCPGGEAPRAKRVLVASCKKRLTRFNOSLKDFTGEARPOKGRK 120

QY 172 PRPRASAKANOAELEWAY 190
DB 121 PRPRASAKANOAELEWAY 139
```

```

RESULT 6
US-09-866-050A-159
; Sequence 159, Application US/09866050A
; Publication No. US20030040471A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Ornust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishnand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 159
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Mouse
US-09-866-050A-159
```

```

Query Match 32.8%; Score 344; DB 9; Length 206;
Best Local Similarity 42.7%; Pred. No. 1.9e-26;
Matches 79; Conservative 30; Mismatches 62; Indels 14; Gaps 8;
```

```

QY 5 AFKNDATETIIPELGYPPELENNKTMRAENGRRPHHPETKDV--EYSCREL 60
DB 23 AFKNDATETI--YSHVVPVPAHPSNSTLNQARNGR--HFSSTGLDRNSRVYGCREL 78

QY 61 HTRVYVTDGPRSAKPYTELVCSCGCGPARLLPNAIGRG--KWM-RPSGDFRCIPRY 116
DB 79 RSTKYSIQGCTSLSPLELVACGCLPLVLPWMTGGGYTKTWSRSGEMCVNDKT 138

QY 117 RAQRVQLCPGGEAPRAKRVLVASCKKRLTRFNOSLKDFTGEARPOKGRKPRRA 176
DB 139 RQRVQLCPGGEAPRAKRVLVASCKKRLTRFNOSLKDFTGEARPOKGRKPRRA 197

QY 177-RSAR 180
DB 198 SKSSK 202
```

```

RESULT 7
US-09-866-050A-286
; Sequence 286, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Ornust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishnand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 286
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Mouse
US-09-866-050A-286
```

```

Query Match 32.8%; Score 344; DB 9; Length 206;
Best Local Similarity 42.7%; Pred. No. 1.9e-26;
Matches 79; Conservative 30; Mismatches 62; Indels 14; Gaps 8;
```

```

QY 5 AFKNDATETIIPELGYPPELENNKTMRAENGRRPHHPETKDV--EYSCREL 60
DB 23 AFKNDATETI--YSHVVPVPAHPSNSTLNQARNGR--HFSSTGLDRNSRVYGCREL 78
```


US-09-867-274-25
 ; Sequence 25, Application US/09867274
 ; Patent No. US20020106650A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Paszty, Christopher
 ; APPLICANT: Gao, Yongming
 ; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
 ; FILE REFERENCE: 01017/37428
 ; CURRENT APPLICATION NUMBER: US/09/867,274
 ; PRIOR FILING DATE: 2001-05-29
 ; PRIOR APPLICATION NUMBER: US 60/208,550
 ; PRIOR FILING DATE: 2000-06-01
 ; PRIOR APPLICATION NUMBER: US 60/223,542
 ; PRIOR FILING DATE: 2000-08-04
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 25
 ; LENGTH: 183
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-867-274-25

Query Match 32.1%; Score 337; DB 10; Length 183;
 Best Local Similarity 42.4%; Pred. No. 7, 9e-26;
 Matches 78; Conservative 30; Mismatches 62; Indels 14; Gaps 8;

QY 6 EKNDATETIPELGEYPP-PELENNKTNRAENGRRPHHPETKDVSEYSCRELH 61
 DB 1 EKNDATETIL--YHVVKVPVPAHPSNSTLQARNGR--HFSNTGLDRNRYVVGRELR 56
 QY 62 FTRYVTDGPRSAKPVTELVCSGCCPARLLPNAIGRG--KMW-RPSGDPFCIPDRYR 117
 DB 57 STYKIDGQCTSTSLPELVACAGECLPVLPMWIGGYTKTWKSRSSQEWRCVWDKTR 116
 QY 118 AGRVOLLCPGGEAPRAKRVLVASCCKRLTRPHNOSELKDFTEARPOKGRKPPRA- 176
 DB 117 TORIQLOCCQG-STRTYKIVTVACKCKRYTRQHNESHNFESMSPAKPVQHRERKRAS 175
 QY 177 RSKAK 180
 DB 176 KSKK 179

RESULT 11
 US-09-814-777A-126
 ; Sequence 126, Application US/09814777A
 ; Patent No. US20020142415A1
 ; GENERAL INFORMATION:
 ; APPLICANT: KOOPMAN, Peter Anthony
 ; APPLICANT: MUSCAT, George Eugene Orlando
 ; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND POLYNUCLEOTIDES AND METHODS OF USING THEM
 ; FILE REFERENCE: 21415-0003
 ; CURRENT APPLICATION NUMBER: US/09/814,777A
 ; PRIOR FILING DATE: 2001-03-23
 ; PRIOR APPLICATION NUMBER: AU P06457
 ; PRIOR FILING DATE: 2000-03-24
 ; NUMBER OF SEQ ID NOS: 128
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 126
 ; LENGTH: 337
 ; TYPE: PRT
 ; ORGANISM: Human
 US-09-814-777A-126

Query Match 9.3%; Score 98; DB 10; Length 337;
 Best Local Similarity 25.8%; Pred. No. 0.063;
 Matches 48; Conservative 15; Mismatches 57; Indels 66; Gaps 10;

QY 13 IIPLEGYPPPEPLENNKTNRAENGRRPHHPETKDVSEYSCRELHTRYVTDG--- 69
 DB 129 LLEPLAP-PPPPPE-----PPPAASGARAFRELPLPLGAFDGLGL 168
 QY 70 PCRSKAEVTELVCSGCCPARLL-PNAIGRKMMRPSGDPFCIPDRYRA-----QRYVOLL 124

DB 169 PTPERSPLDGL-----EPGEAFAFPPEPAPRTARWRPSAP-----PTAHRVVGCPRLILRG 219
 QY 125 CPGG-----EAPRAKRVLVAS-----CKKRLTRFNHQSLEKDGCTAARPOKGRK 171
 DB 220 SPGGAGODRAPRASLACTTAPMARPARTPARCRR-----RPRFRWRA 263
 QY 172 PPRPAR 177
 DB 264 PPGAR 269

RESULT 12
 US-10-044-716-10
 ; Sequence 10, Application US/10044716
 ; Patent No. US20020159986A1
 ; GENERAL INFORMATION:
 ; APPLICANT: LANGENFELD, John
 ; TITLE OF INVENTION: BONE MORPHOGENETIC PROTEIN-2 IN THE TREATMENT AND DIAGNOSIS OF CANCER
 ; FILE REFERENCE: 270/070US
 ; CURRENT APPLICATION NUMBER: US/10/044,716
 ; PRIOR FILING DATE: 2002-01-11
 ; PRIOR APPLICATION NUMBER: US60/261,252
 ; PRIOR FILING DATE: 2001-01-12
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 10
 ; LENGTH: 184
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-044-716-10

Query Match 8.5%; Score 89.5; DB 9; Length 184;
 Best Local Similarity 24.1%; Pred. No. 0.021; Mismatches 57; Indels 41; Gaps 9;

QY 18 GEYPPPELENNK-----TNRAENGRRPHHPETKDVSEYSCRELHFT--R 64
 DB 31 GAIP-PPKAGHNDSEQTPQPGSRNRGRGQGTAMPGE--EVLESQDALHVTERRK 87
 QY 65 YTDGPCR-----SAKPVTELVCSGCCPARLLPNAIGRKMMRPSGDPDR- 110
 DB 88 YLRDWCCTQPLKQTHIEGCGNSRTIINRCYGCQN-SFYIPRH-----RKEGSPQS 140
 QY 111 ---CIPDRYRAQRYVOLLCPGGEAP-RARKRVLVASCCK 144
 DB 141 CSFCKPKKFTTMMVTLCBLOPFTKGRVTRVKOCRC 178

RESULT 13
 US-10-159-749-2
 ; Sequence 2, Application US/10159749
 ; Publication No. US20020192219A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harland, Richard
 ; APPLICANT: Hsu, David
 ; TITLE OF INVENTION: Morphogenic Proteins
 ; NUMBER OF SEQUENCES: 13
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP
 ; STREET: 75 DENISE DRIVE
 ; CITY: HILLSBOROUGH
 ; STATE: CALIFORNIA
 ; COUNTRY: USA
 ; ZIP: 94010
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/159,749
 ; FILING DATE: 29-May-2002

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/040,229
FILING DATE: 13-Mar-1998
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B97-020-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 184 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-159-749-2

Query Match 8.5%; Score 89.5; DB 9; Length 184;
Best Local Similarity 24.1%; Pred. No. 0.21;
Matches 38; Conservative 22; Mismatches 57; Indels 41; Gaps 9;

QY 18 GEYEPPELENNK-----TNRAENGRRPHHPETKDVSEYSCRELHFT--R 64
DB 31 GAIP-PPKAGHNDSEQTSPQSPGSRNRGRGTAMPGE--EVLSSQELHVTERR 87

QY 65 YVTGDPGR-----SAKPYTELVSQGGCPARLLPNAIGKMMRBSGPPFR- 110
DB 88 YLRKDWCKTQPLKQTHIEGNSRTIINRFYGCN-SFYIPRH-----RKEGSGFQS 140

QY 111 --CIPRYRAQRVQLCPGGEAP-RARKVRLVASCRC 144
DB 141 CSFCKPKKFTTMMVTLNCPELQPRKKKRVTRVACRC 178

RESULT 14
US-10-159-749-6
Sequence 6, Application US/10159749
Publication No. US20020192219A1
GENERAL INFORMATION:
APPLICANT: Harland, Richard
Hsu, David
TITLE OF INVENTION: Morphogenic Proteins
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/159,749
FILING DATE: 29-May-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/040,229
FILING DATE: 13-Mar-1998
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B97-020-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:
LENGTH: 184 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-159-749-6

Query Match 8.3%; Score 87.5; DB 9; Length 184;
Best Local Similarity 22.4%; Pred. No. 0.32;
Matches 34; Conservative 24; Mismatches 55; Indels 39; Gaps 8;

QY 23 PPELE-----NNKTNRAENGRRPHHPETKDVSEYSCRELHFT--RYTDP 70
DB 36 PPDDKQPNDSQWQOQSGSRHREKGTSMPE--EVLSSQELHVTERRKLDW 93

QY 71 CR-----SAKPYTELVSQGGCPARLLPNAIGKMMRBSGPPFR----CIP 113
DB 94 CKTQPLKQTHIEGNSRTIINRFYGCN-SFYIPRHV-----RKEGSGFQSCFCKP 146

QY 114 DRYRAQRVQLCPGGEAPRARK-VRLVASCRC 144
DB 147 KFTTMMVTLNCPELQPRKKKRVTRVACRC 178

RESULT 15
US-10-159-749-4
Sequence 4, Application US/10159749
Publication No. US20020192219A1
GENERAL INFORMATION:
APPLICANT: Harland, Richard
Hsu, David
TITLE OF INVENTION: Morphogenic Proteins
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/159,749
FILING DATE: 29-May-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/040,229
FILING DATE: 13-Mar-1998
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B97-020-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 182 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-159-749-4

Query Match 8.2%; Score 86.5; DB 9; Length 182;
Best Local Similarity 22.8%; Pred. No. 0.4;
Matches 34; Conservative 24; Mismatches 64; Indels 27; Gaps 7;

```

OY 18 GEYBEPPELENNKTNRAENG---GRPPHHPFETKDVSEYSCRELAHT--RYTDGPC 71
Db 33 GAIPPDKGOPNDSEOGQAOPEGDRVRGKGQALAEVLESQELHTTERKYLKRDWC 92
OY 72 R-----SAKPYTELVCSSQCCPARLLPNAIGR--GKWTBPSGDPFRCIPDRY 116
Db 93 KTOPLKQTHEDGNSRTIINRFCYQCN-SFYIPRHIREEGSFQCS---FCRPKKF 147
OY 117 RAORVOLCPGGEAP--RARVRVLVASCRC 144
Db 148 TTMVVTLNCBELQPTTKKRIITRVKQCR 176

```

Search completed: March 28, 2003, 14:20:45
 Job time : 14.16 secs

XX (AMGE-) AMGEN INC.
 PA Paszty CJ, Gao Y;
 PI WPI; 2002-114325/15.
 XX DR N-PSDB; ABA94293.
 XX PT New human and mouse cysteine-knot polypeptide designated as Cloaked-2,
 PT for treating or preventing kidney, heart (e.g. myocardial infarction)
 PT or liver (e.g. hepatitis) diseases
 XX
 PS Claim 13; Fig 1; 170pp; English.
 XX
 CC The invention relates to polypeptides comprising a cysteine knot motif
 CC and designated as Cloaked-2, derived from human and mouse. The cloaked-2
 CC polypeptides can be expressed by standard recombinant methodology. The
 CC cloaked-2 polynucleotides are useful in gene therapy and antisense
 CC therapy. The cloaked-2 polypeptides and polynucleotides are useful for
 CC treating, preventing, ameliorating or detecting diseases and disorders of
 CC the kidney (e.g. anemia, hypertension or low blood pressure), heart (e.g.
 CC cardiac hypertrophy, congestive heart failure, myocardial infarction,
 CC arrhythmias, atherosclerosis, hypertension or low blood pressure),
 CC skeletal muscle (e.g. muscular dystrophy or cachexia), placenta (e.g.
 CC congenital abnormalities or miscarriage), liver (e.g. hepatitis or
 CC cirrhosis), pancreas (e.g. diabetes or pancreatitis), thyroid (e.g.
 CC Graves' disease or myxedema) or adrenal cortex (e.g. Cushing's disease
 CC or Addison's disease), homeostasis or metabolic diseases (e.g. obesity,
 CC cancer or myopathies), infections, or autoimmune diseases. Selective
 CC binding agents may be used to modulate the biological activities of
 CC Cloaked-2 polypeptides or to detect Cloaked-2 polypeptide levels in a
 CC sample. Transgenic non-human animals are useful for drug candidate
 CC screening. The present sequence represents the human cloaked-2
 CC mature polypeptide sequence.

XX Sequence 190 AA;

Query Match 100.0%; Score 190; DB 23; Length 190;

Best Local Similarity 100.0%; Pred. No. 1.9e-190; Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGWAQFNKDAETIIPELGEYEPPELENKTMRAENGGRPPHPPETKDVSEYSREL 60
 DB 1 QGWAQFNKDAETIIPELGEYEPPELENKTMRAENGGRPPHPPETKDVSEYSREL 60
 QY 61 HFTRYVTDDGPRSAKPTVELVCSGCGPARLLPNAIGRKWMRSGDFRCIPDRYRAQR 120
 DB 61 HFTRYVTDDGPRSAKPTVELVCSGCGPARLLPNAIGRKWMRSGDFRCIPDRYRAQR 120
 QY 121 VOLLCPGSEAPRAKRVLVASCKCKRLTRFNOSSELDQFTGARPPQKGRKPRPARSAK 180
 DB 121 VOLLCPGSEAPRAKRVLVASCKCKRLTRFNOSSELDQFTGARPPQKGRKPRPARSAK 180
 QY 181 ANQAELENAY 190
 DB 181 ANQAELENAY 190

RESULT 2

AAB26106 AAB26106 standard; Protein; 213 AA.

XX AAB26106;

XX 15-JAN-2001 (first entry)

XX Human DAN/Cerberus-related protein 6 (hDCR6) (exons 1 and 4).

XX Human; DNA/Cerberus-related protein 6; hDCR6; morphogenic protein;
 KW antagonist; BMP; cell growth; cell differentiation; bone formation;
 KM gene therapy.
 XX
 OS Homo sapiens.

XX W020005193-A2.
 PN 21-SEP-2000.
 XX
 PD 02-MAR-2000; 2000WO-US05537.
 XX
 PF 12-MAR-1999; 99US-0124118.
 XX
 PR (REG-) REGENERON PHARM INC.
 XX
 PA Economides AN;
 XX
 PI WPI; 2000-638179/61.
 DR N-PSDB; AAA94051.
 XX
 CC Novel isolated, human DNA/Cerberus related protein 6 which include
 CC natural homologue, and polypeptides comprising DCR6 domain and nucleic
 CC acids encoding the proteins which are useful as probes and primers
 CC
 PS Claim 8; Fig 3; 40pp; English.
 XX
 CC The present sequence comprises the amino acid sequence encoded by exons 1
 CC and 4 of the human DAN/Cerberus-related protein 6 (hDCR6) coding
 CC sequence. The coding sequence was isolated from a human kidney cDNA
 CC library containing exons 1 and 4 of the sequence. hDCR6 is closely
 CC related to the DAN and DCR5 proteins, both of which act as antagonists of
 CC morphogenic proteins such as BMP. It is possible that the hDCR6 gene and
 CC protein can be used as immunogens, modulators of cell function, growth
 CC and differentiation, to reduce undesirable bone formation, to identify
 CC DCR6 binding agents, in diagnosis, and in gene therapy.

XX Sequence 213 AA;

Query Match 100.0%; Score 190; DB 21; Length 213;

Best Local Similarity 100.0%; Pred. No. 2.1e-190; Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGWAQFNKDAETIIPELGEYEPPELENKTMRAENGGRPPHPPETKDVSEYSREL 60
 DB 24 QGWAQFNKDAETIIPELGEYEPPELENKTMRAENGGRPPHPPETKDVSEYSREL 83
 QY 61 HFTRYVTDDGPRSAKPTVELVCSGCGPARLLPNAIGRKWMRSGDFRCIPDRYRAQR 120
 DB 84 HFTRYVTDDGPRSAKPTVELVCSGCGPARLLPNAIGRKWMRSGDFRCIPDRYRAQR 143
 QY 121 VOLLCPGSEAPRAKRVLVASCKCKRLTRFNOSSELDQFTGARPPQKGRKPRPARSAK 180
 DB 144 VOLLCPGSEAPRAKRVLVASCKCKRLTRFNOSSELDQFTGARPPQKGRKPRPARSAK 203
 QY 181 ANQAELENAY 190
 DB 204 ANQAELENAY 213

RESULT 3

AA96429 AA96429 standard; Protein; 213 AA.

XX AA96429;

XX 12-SEP-2000 (first entry)

XX Human TGF-beta binding protein (BERR).

XX osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KW BERR; gene therapy; antisense therapy; fracture; bone mineralization.
 XX
 OS Homo sapiens.

XX WO200032773-A1.
 PN 08-JUN-2000.

[illegible]

XX	PN	WO200075317-A2.
XX	PD	14-DEC-2000.
XX	PF	15-MAY-2000; 2000WO-US13358.
XX	PR	09-JUN-1999; 99US-0138385.
XX	PR	20-JUL-1999; 99US-0144790.
XX	PR	03-AUG-1999; 99US-0146843.
XX	PR	10-AUG-1999; 99US-0148188.
XX	PR	17-AUG-1999; 99US-0149320.
XX	PR	17-AUG-1999; 99US-0149327.
XX	PR	17-AUG-1999; 99US-0149396.
XX	PR	20-AUG-1999; 99US-0150114.
XX	PR	31-AUG-1999; 99US-0151700.
XX	PR	31-AUG-1999; 99US-0151734.
XX	PA	(GETH) GENENTECH INC.
XX	PI	Botstein DA, Goddard A, Gurney AL, Smith V, Watanabe CK, Wood WI;
XX	DR	WPI: 2001-071075/108.
XX	DR	N-PSDB; AAA91023.
XX	PT	Antibodies against PRO polypeptides, useful for diagnosing and treating
XX	PT	tumours are associated with gene amplification, neoplastic cell growth
XX	PT	and proliferation in mammals -
XX	PS	Claim 61; Fig 20; 143pp; English.
XX	CC	This sequence is a human PRO protein of the invention. The PRO
XX	CC	proteins are secreted proteins. Antagonists or antibodies of PRO
XX	CC	polypeptides are useful for diagnosing and treating tumours are
XX	CC	associated with gene amplification, neoplastic cell growth and
XX	CC	proliferation in mammals, and those conditions characterised by
XX	CC	overexpression and/or activation of the amplified genes. Such conditions
XX	CC	include benign or malignant tumours (e.g. renal, liver, kidney, bladder,
XX	CC	breast, gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,
XX	CC	thyroid, hepatic carcinomas, sarcomas, glioblastomas and various head and
XX	CC	neck tumours); leukemias and lymphoid malignancies; neuronal, glial,
XX	CC	astrocytal, hypothalamic, and other glandular, macrophageal, epithelial,
XX	CC	stromal and blastocoelec disorders; and inflammatory, angiogenic and
XX	CC	immunologic disorders. These may further be used to qualitatively or
XX	CC	quantitatively detect the expression of proteins encoded by the
XX	CC	amplified genes, and in tumour diagnostics or prognostics. The PRO
XX	CC	polypeptide or its antagonist may be used for the preparation of a
XX	CC	medicament in the treatment of a condition, which is responsive to the
XX	CC	PRO polypeptide, its antagonist or anti-PRO antibody.
XX	SO	Sequence 213 AA;
QY		Query Match 100.0%; Score 190; DB 22; Length 213;
QY		Best Local Similarity 100.0%; Pred. No. 2.1e-190;
QY		Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db		1 QGMAQFKNDATREIIPELGEPYEPPELENNNTNMAENGSRPHNHPRTKVSYSCREL 60
QY		24 QGMAQFKNDATREIIPELGEPYEPPELENNNTNMAENGSRPHNHPRTKVSYSCREL 83
Db		61 HFTRTVTDGPRSAKPVETLVCSSGCGPARLLPNAIGRGKWRPESGDPFRCTPDRYRAQR 120
QY		84 HFTRTVTDGPRSAKPVETLVCSSGCGPARLLPNAIGRGKWRPESGDPFRCTPDRYRAQR 143
Db		121 VOLLCPGGEAPRAKRVRLVASCCKRLTRFTNOSLELKGFTDARPOGGRKRRPARSAK 180
QY		144 VOLLCPGGEAPRAKRVRLVASCCKRLTRFTNOSLELKGFTDARPOGGRKRRPARSAK 203
Db		181 ANOAELENAVY 190
QY		204 ANOAELENAVY 213

RESULT 5
 ABG34061
 ID ABG34061 standard; Protein; 213 AA.
 AC ABG34061;
 DT 15-JUL-2002 (first entry)
 DE Human Pro peptide #32.
 KW Human; PRO; secreted protein; transmembrane protein;
 genetic disorder; tumour; cancer.
 OS Homo sapiens.
 PN W0200224888-A2.
 PD 28-MAR-2002.
 PF 29-AUG-2001; 2001WO-US27099.
 PR 01-SEP-2000; 2000US-229896P.
 PR 05-SEP-2000; 2000US-230621P.
 PR 22-SEP-2000; 2000US-235147P.
 PR 10-NOV-2000; 2000WO-US30873.
 PR 12-JAN-2001; 2001US-261878P.
 PR 16-JAN-2001; 2001US-261910P.
 PR 16-JAN-2001; 2001US-261939P.
 PR 16-JAN-2001; 2001US-262150P.
 PR 25-JAN-2001; 2001US-264395P.
 PR 02-FEB-2001; 2001US-266421P.
 PR 09-FEB-2001; 2001US-267623P.
 PR 28-FEB-2001; 2001WO-US06520.
 PR 09-MAR-2001; 2001US-274399P.
 PR 03-APR-2001; 2001US-280982P.
 PR 04-APR-2001; 2001US-282129P.
 PR 04-APR-2001; 2001US-282199P.
 PR 09-MAY-2001; 2001US-290589P.
 PR 25-MAY-2001; 2001WO-US17092.
 PR 01-JUN-2001; 2001WO-US17800.
 PR 20-JUN-2001; 2001WO-US19692.
 PR 29-JUN-2001; 2001WO-US21066.
 PR 09-JUL-2001; 2001WO-US21735.
 XX
 XX (GENTH) GENENTECH INC.
 PA Baker KP, Baton DL, Filvaroff E, Goddard A, Grimaldi JC,
 PI Gurney AL, Smith V, Stephan J, Watanabe CK, Wood WL, Zhang Z,
 PI Fong S;
 XX WPI; 2002-362426/39.
 DR N-PSDB; ABK69992.
 PT New PRO polypeptides and polynucleotides encoding the polypeptides,
 PT useful in gene therapy, chromosome identification, tissue typing, or
 PT for genetic analysis of individuals with genetic disorders -
 XX
 PS Claim 11, Figure 64; 218pp; English.

CC PRO polypeptides are useful in gene therapy, and as molecular weight
 CC markers for protein electrophoresis purposes. The sequences may
 CC also be used to detect overexpression on PRO polypeptides in cancerous
 CC tumours and for screening for differentially expressed genes using
 CC microarray technology. The present sequence represents a human PRO
 CC protein of the invention.
 XX
 SQ Sequence 213 AA;
 Query Match 100.0%; Score 190; DB 23; Length 213;
 Best Local Similarity 100.0%; Pred. No. 2, 1e-190;
 Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 OGMQAFKNDATETIIPETIGEYEPPEPELENNKTKMRAENGGRPHHPETDVSYSREL 60
 DB 24 OGMQAFKNDATETIIPETIGEYEPPEPELENNKTKMRAENGGRPHHPETDVSYSREL 83
 QY 61 HFTRYVTDGPCRSAPVTELVCSGGCGPARLLPNAIGRKMRSPGDFRCIPRYAQR 120
 DB 84 HFTRYVTDGPCRSAPVTELVCSGGCGPARLLPNAIGRKMRSPGDFRCIPRYAQR 143
 QY 121 VOLLCPGGAPRARKVRLVASCCKRLTRFNQSELDGFTGARPPQKGRKPRPARSAK 180
 DB 144 VOLLCPGGAPRARKVRLVASCCKRLTRFNQSELDGFTGARPPQKGRKPRPARSAK 203
 QY 161 ANQAELENAY 190
 DB 204 ANQAELENAY 213
 XX
 XX RESULT 6
 XX ABB07209
 XX ID ABB07209 standard; Protein; 213 AA.
 XX AC ABB07209;
 XX DT 26-MAR-2002 (first entry)
 XX DE Human cloaked-2 polypeptide sequence.
 XX
 XX Cloaked-2; cysteine knot motif; nephrotropic; cardiac; immunomodulator;
 XX hepatotropic; antiinflammatory; antihypertoid; cytostatic; neuroprotective;
 XX antianemic; hypotensive; antiarrhythmic; antiarteriosclerotic; muscular;
 XX antidiabetic; anorectic; gene therapy; cell therapy; antisense therapy;
 XX human.
 XX
 XX Homo sapiens.
 XX OS
 XX FH
 XX Key Location/Qualifiers
 XX Peptide 1..23
 XX FT /note= "signal peptide"
 XX FT Protein 24..213
 XX FT /note= "mature protein (ABB07207)"
 XX
 XX W0200192308-A2.
 XX PD 06-DEC-2001.
 XX PF 29-MAY-2001; 2001WO-US17478.
 XX PR 01-JUN-2000; 2000US-208550P.
 XX PR 04-AUG-2000; 2000US-223542P.
 XX PA (AMGE-) AMGEN INC.
 XX PI Paszty CJ, Gao Y;
 XX WPI; 2002-114325/15.
 DR N-PSDB; ABA94293.
 XX
 XX New human and mouse cysteine-knot polypeptide designated as Cloaked-2,
 PT for treating or preventing kidney, heart (e.g. myocardial infarction)
 PT or liver (e.g. hepatitis) diseases -

DE Human TGF-beta binding protein (BEER) variant V101.

XX osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KM BEER; variant; V101; gene therapy; antisense therapy; fracture;
 KW Chromosome 17q12-21; bone mineralization.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 10

FT /label= V101

FT /note= "wild type valine has been substituted with
 isoleucine"

XX MO200032773-A1.

XX 08-JUN-2000.

XX 24-NOV-1999; 99WO-US27990.

XX 27-NOV-1998; 98US-0110283.

XX (DARW-) DARWIN DISCOVERY LTD.

XX Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepier BW;

XX Van Ness J, Winkler DG;

XX WPI; 2000-412321/35.

XX N-PSDB; AAA29056.

XX Nucleic acids (1) encoding a transforming growth factor beta binding

PT protein, useful for identifying agents for treating osteopenia,

PT osteoporosis and fractures

XX Claim 3; Page 119-120; 162pp; English.

XX This shows a variant human transforming growth factor-beta (TGF-beta)
 CC binding protein designated BEER V101, which comprises a substitution of
 CC isoleucine for the wild-type valine at residue 10. The cDNA and protein
 CC may be used for prevention, treatment and diagnosis of diseases
 CC associated with inappropriate BEER expression. For example, they may be
 CC used to treat disorders associated with decreased TGF-beta BP expression.
 CC The cDNA or vectors may be administered to treat diseases by rectifying
 CC mutations or deletions in a patient's genome that affect the activity of
 CC BEER by expressing inactive proteins or to supplement the patients own
 CC production of BEER polypeptides. The nucleic acids may be used for
 CC recombinant production of BEER, gene therapy, antisense therapy, as
 CC probes for diagnostic assays and for functional studies. BEER may be used
 CC to raise antibodies and for identification of BEER modulators. BEER
 CC antagonists may be used to increase bone mineral content for the
 CC treatment of disorders such as osteopenia, osteoporosis, fractures and
 CC other disorders associated with low mineral content.

XX Sequence 213 AA;

Query Match 92.1%; Score 175; DB 21; Length 213;

Best Local Similarity 100.0%; Pred. No. 1.1e-174;

Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ELGEPPEPELENNKTMNRAENGRRPPHPPETKDVSEVSCRELHTRVYTDGPPCSAK 75

DB 39 ELGEPPEPELENNKTMNRAENGRRPPHPPETKDVSEVSCRELHTRVYTDGPPCSAK 98

QY 76 PVTLEVSGGCGPARLLPNAIGRGKMWPSGDPDRCTPDYRAQVOLLCPGSEAPPAK 135

DB 99 PVTLEVSGGCGPARLLPNAIGRGKMWPSGDPDRCTPDYRAQVOLLCPGSEAPPAK 158

QY 136 VRLVASKCKRLTFPHNOSELKDGCTEAPROKRRPRPARSKANOALENNAY 190

DB 159 VRLVASKCKRLTFPHNOSELKDGCTEAPROKRRPRPARSKANOALENNAY 213

RESULT 9

AAV96436

ID AAV96436 standard; Protein; 213 AA.

XX AAV96436;

XX 12-SEP-2000 (first entry)

DE Human TGF-beta binding protein (BEER) variant P38R.

XX osteopathic; transforming growth factor-beta; TGF-beta; binding protein;

KM BEER; variant; P38R; gene therapy; antisense therapy; fracture;

KW Chromosome 17q12-21; bone mineralization.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 38

FT /label= P38R

FT /note= "wild type proline has been substituted with
 arginine"

XX MO200032773-A1.

XX 08-JUN-2000.

XX 24-NOV-1999; 99WO-US27990.

XX 27-NOV-1998; 98US-0110283.

XX (DARW-) DARWIN DISCOVERY LTD.

XX Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepier BW;

XX Van Ness J, Winkler DG;

XX WPI; 2000-412321/35.

XX N-PSDB; AAA29062.

XX Nucleic acids (1) encoding a transforming growth factor beta binding

PT protein, useful for identifying agents for treating osteopenia,

PT osteoporosis and fractures

XX Disclosure; Page 121; 162pp; English.

XX This shows a variant human transforming growth factor-beta
 CC (TGF-beta) binding protein designated BEER P38R. The encoded protein
 CC comprises a substitution of arginine for the wild-type proline at
 CC residue 38. The cDNA and protein may be used for prevention, treatment
 CC and diagnosis of diseases associated with inappropriate BEER expression.
 CC For example, they may be used to treat disorders associated with
 CC decreased TGF-beta BP expression. The cDNA or vectors may be administered
 CC to treat diseases by rectifying mutations or deletions in a patient's
 CC genome that affect the activity of BEER by expressing inactive proteins
 CC or to supplement the patients own production of BEER polypeptides. The
 CC nucleic acids may be used for recombinant production of BEER, gene
 CC therapy, antisense therapy, as probes for diagnostic assays and for
 CC functional studies. BEER may be used to raise antibodies and for
 CC identification of BEER modulators. BEER antagonists may be used to
 CC increase bone mineral content for the treatment of disorders such as
 CC osteopenia, osteoporosis, fractures and other disorders associated with
 CC low mineral content.

XX Sequence 213 AA;

Query Match 92.1%; Score 175; DB 21; Length 213;

Best Local Similarity 100.0%; Pred. No. 1.1e-174;

Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ELGEPPEPELENNKTMNRAENGRRPPHPPETKDVSEVSCRELHTRVYTDGPPCSAK 75

DB 39 ELGEPPEPELENNKTMNRAENGRRPPHPPETKDVSEVSCRELHTRVYTDGPPCSAK 98

QY 76 PVTLEVSGGCGPARLLPNAIGRGKMWPSGDPDRCTPDYRAQVOLLCPGSEAPPAK 135

Qy 375 CTTGCTGCCCCAAGCCATCGGCGCGGCAAGTGTGCGA CTTAGTGGCCCGACTTCCG 434
 Db 3206 CTTGCTGCCCCAAGCCATCGGCGCGGCAAGTGTGCGA CTTAGTGGCCCGACTTCCG 3265
 Qy 435 CTGATATCCCCGACCGCTACCGCGCGGCGGCAAGTGTGCGA CTTAGTGGCCCGACTTCCG 494
 Db 3266 CTGATATCCCCGACCGCTACCGCGCGGCGGCAAGTGTGCGA CTTAGTGGCCCGACTTCCG 3325
 Qy 495 GCGCGCGCGGCGGCAAGTGTGCGA CTTAGTGGCCCGACTTCCG 554
 Db 3326 GCGCGCGCGGCGGCAAGTGTGCGA CTTAGTGGCCCGACTTCCG 3385
 Qy 555 CCACACCAAGTGTGCGA CTTAGTGGCCCGACTTCCG 614
 Db 3386 CCACACCAAGTGTGCGA CTTAGTGGCCCGACTTCCG 3445
 Qy 615 GAAGCGCGCGGCGGCAAGTGTGCGA CTTAGTGGCCCGACTTCCG 674
 Db 3446 GAAGCGCGGCGGCGGCAAGTGTGCGA CTTAGTGGCCCGACTTCCG 3505
 Qy 675 CTAGAGCGCGGCGGCGGCAAGTGTGCGA CTTAGTGGCCCGACTTCCG 734
 Db 3506 CTAGAGCGCGGCGGCGGCAAGTGTGCGA CTTAGTGGCCCGACTTCCG 3565
 Qy 735 CATTCTGTCTCTGCGCGGTGTT 759
 Db 3566 CATTCTGTCTCTGCGCGGTGTT 3590

RESULT 11
 AAD27577/c
 ID AAD27577 standard; DNA; 7099 BP.

XX AAD27577;

XX 18-APR-2002 (first entry)

XX Human osteolevin gene.

XX Human; osteolevin; osteopathic; cyrostatic; bone formation; osteoporosis;
 KM Van Buchem-sclerosteosis disease; sclerosteosis; transgenic animal; gene;
 KM Paget's disease; Chromosome 17; ds.

XX Homo sapiens.

XX Key Location/Qualifiers
 FH poly_a_site complement (315..503)
 FT complement (503..1624)

FT 3'UTR complement (503..1624)
 FT complement (503..1624)

FT variation replace (525, C)
 FT complement (525, C)

FT variation replace (916, T)
 FT complement (916, T)

FT variation replace (1054, A)
 FT complement (1054, A)

FT variation replace (1358, T)
 FT complement (1358, T)

FT exon complement (1920..2341)
 FT complement (1920..2341)

FT CDS complement (1920..5319)
 FT complement (1920..5319)

FT variation replace (2186, A)
 FT complement (2186, A)

FT variation replace (2186, A)
 FT complement (2186, A)

FT variation replace (2186, A)
 FT complement (2186, A)

FT variation replace (2186, A)
 FT complement (2186, A)

FT variation replace (2186, A)
 FT complement (2186, A)

FT variation replace (2186, A)
 FT complement (2186, A)

FT variation replace (2186, A)
 FT complement (2186, A)

FT variation /*tag= k
 FT replace (3539, T)
 FT /*tag= l
 FT replace (3944, T)
 FT variation /*tag= m
 FT replace (4425, G)
 FT variation /*tag= n
 FT replace (5097, A)
 FT intron /*tag= o
 FT complement (5099..2242)
 FT /*tag= p
 FT complement (5100..5331)
 FT /*tag= q
 FT /number= 1
 FT replace (5292, T)
 FT /*tag= r
 FT complement (5366..5405)
 FT /*tag= s
 FT replace (5666, A)
 FT variation /*tag= t
 FT replace (5696, T)
 FT variation /*tag= u
 FT replace (5833, C)
 FT variation /*tag= v
 FT replace (6070, G)
 FT variation /*tag= w
 FT replace (6392, G)
 FT variation /*tag= x
 FT replace (6474, G)
 FT variation /*tag= y
 FT replace (6718..6719, CTCCT)
 FT variation /*tag= z
 FT replace (6737, A)
 FT /*tag= aa
 FT replace (6867, C)
 FT variation /*tag= ab
 FT replace (6926, G)
 FT /*tag= ac
 FT replace (6927, T)
 FT variation /*tag= ad

FN WO200198491-A2.

PD 27-DEC-2001.

PF 15-JUN-2001; 2001WO-EP06795.

PR 19-JUN-2000; 2000EP-0112867.

PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
 PA (UYIN-) UNIV INSTELLING ANTWERPEN UTA.

PI Bailemans W, Ebeling M, Foernzler D, Patel N, Van Hul W;
 PI Vickery BH;

DR WPI; 2002-139789/18.

DR P-PSDB; AAEL7089.

PT Novel genetic polymorphisms in the Van Buchem-sclerosteosis disease
 PT region that are associated with abnormal bone formation useful for
 PT diagnosis and assessment of osteoporosis or sclerosteosis in humans
 XX

Claim 11; Fig 3; 70pp; English.

XX The invention relates to a nucleic acid encoding osteolevin region
 CC polymorphisms. The invention also relates to genetic polymorphisms in
 CC the Van Buchem-sclerosteosis disease region that are associated with
 CC disorders resulting in either net excess bone formation or insufficient
 CC bone formation in humans. Osteolevin DNA is useful for screening for
 CC osteolevin polymorphisms associated with abnormal bone formation in a
 CC subject and for the presence of a heritably linked form of abnormal bone
 CC formation in a subject, by determining the presence of a polymorphism in
 CC the osteolevin nucleic acid sequence obtained from the subject.

QY 555 CCACACAGTGGAGCTCAAGACTTCGAGACCGAGCCGCTCGGCCGAGAAAGGCG 614
 DB 3485 CCACACAGTGGAGCTCAAGACTTCGAGACCGAGCCGCTCGGCCGAGAAAGGCG 3544
 QY 615 GAAGCCCGCGCCCGCGCCGCGGAGCGCCAAAGCCACAGGCGAGCTGAGAAACGCTTA 674
 DB 3545 GAAGCCCGCGCCCGCGCCGCGGAGCGCCAAAGCCACAGGCGAGCTGAGAAACGCTTA 3604
 QY 675 CTAGAGCCCGCGCCCGCGCCCTCCACACCGGCGGCGCCCGCCCTGAAACCGCGCCCA 734
 DB 3605 CTAGAGCCCGCGCCCGCGCCCTCCACACCGGCGGCGCCCGCCCTGAAACCGCGCCCA 3664
 QY 735 CATTTCTGCTGCTGCGCGTGTGTT 759
 DB 3665 CATTTCTGCTGCTGCGCGTGTGTT 3689

RESULT 13

AAK4509/c
 ID AAK45509 standard; DNA; 422 BP.

AC AAK45509;

DT 06-NOV-2001 (first entry)

XX Human bone marrow expressed single exon probe SEQ ID NO: 20066.

KW Human; bone marrow expressed exon; gene expression analysis; probe;

KW microarray; cancer; leukemia; lymphoma; myeloma; ss.

OS Homo sapiens.

PN NO200157276-A2.

XX 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00668.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human bone marrow -

XX Example 4; SEQ ID NO: 20066; 658bp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

XX probes which are derived from genomic sequences expressed in the human

XX bone marrow. They can be used to measure gene expression in bone marrow

XX samples, which may enable the improved diagnosis and treatment of cancers

XX such as lymphoma, leukemia and myeloma. The present sequence is one of

XX the probes of the invention.

XX Sequence 422 BP; 53 A; 145 C; 159 G; 65 T; 0 other;

XX Query Match 55.6%; Score 422; DB 22; Length 422;

XX Best Local Similarity 100.0%; Pred. No. 3e-166;

XX Matches 422; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 257 ACGTGTCCGAGTACGAGTCCGCGAGCTGCACTTCAACCGCTACGTAGCGATGCGCGT 316

DB 422 ACGTGTCCGAGTACGAGTCCGCGAGCTGCACTTCAACCGCTACGTAGCGATGCGCGT 363
 QY 317 GCCGAGGCGCCAAAGCCGCTCAACGAGCTGAGTGTCTCCGCGCAAGTGCGGCCGCGGCC 376
 DB 362 GCCGAGGCGCCAAAGCCGCTCAACGAGCTGAGTGTCTCCGCGCAAGTGCGGCCGCGGCC 303
 QY 377 TGCTGCCCAAGCCGCTACCGCGCGCGCAAGTGTGTGCGCACTAGTGAGGCCGCTTCGCT 436
 DB 302 TGCTGCCCAAGCCGCTACCGCGCGCGCAAGTGTGTGCGCACTAGTGAGGCCGCTTCGCT 243
 QY 437 GCATCCCGGACCGCTACCGCGCGCGCAAGTGTGTGCGCACTAGTGAGGCCGCTTCGCT 496
 DB 242 GCATCCCGGACCGCTACCGCGCGCGCAAGTGTGTGCGCACTAGTGAGGCCGCTTCGCT 183
 QY 497 CGCGGCGCGCAAGTGTGTGCGCTGAGTGTGCGCAAGTGTGCGCAAGTGTGCGCTTCGCT 556
 DB 182 CGCGGCGCGCAAGTGTGTGCGCTGAGTGTGCGCAAGTGTGCGCAAGTGTGCGCTTCGCT 123
 QY 557 ACAACCAAGTCCGAGCTCAAGACTTCGAGACCGAGCCGCTCGGCCGAGAAAGGCGCGGA 616
 DB 122 ACAACCAAGTCCGAGCTCAAGACTTCGAGACCGAGCCGCTCGGCCGAGAAAGGCGCGGA 63
 QY 617 AGCCGCGCGCCCGCGCGCGCGCGCGCGCAAGCCCAACCAAGCCGAGCTGAGAAAGGCGCTACT 676
 DB 62 AGCCGCGCGCCCGCGCGCGCGCGCGCGCGCAAGCCCAACCAAGCCGAGCTGAGAAAGGCGCTACT 3
 QY 677 AG 678
 DB 2 AG 1

RESULT 14

ABS19777/c
 ID ABS19777 standard; DNA; 422 BP.

XX ABS19777;

DT 19-AUG-2002 (first entry)

DE Human genome-derived single exon ORF from lung SEQ ID No 19768.

KW Human; ds, single exon probe; asthma; lung cancer; COPD; IID;

KW chronic obstructive pulmonary disease; interstitial lung disease;

KW familial idiopathic pulmonary fibrosis; neurofibromatosis;

KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;

KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary hemosiderosis;

KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;

KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;

KW primary ciliary dyskinesia; pulmonary hypertension;

KW hyaline membrane disease; open reading frame; ORF.

XX Homo sapiens.

XX WO200186003-A2.

XX 15-NOV-2001.

XX 30-JAN-2001; 2001WO-US00665.

XX 04-FEB-2000; 2000US-180312P.

XX 26-MAY-2000; 2000US-207456P.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-234687P.

XX 27-SEP-2000; 2000US-236359P.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2002-114183/15.

PT Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples -

PS Claim 4; SEQ ID No 19768; 634bp; English.

CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of
CC probes; the novel set of probes which hybridise at high stringency to a
CC nucleic acid expressed in the human lung; measuring gene expression in a
CC sample derived from human lung, comprising (a) contacting the array with
CC a collection of detectably labeled nucleic acid derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of
CC the array; identifying exons in a eukaryotic genome, comprising
CC (a) algorithmically predicting at least one exon from genomic sequences
CC of the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene
CC expression analysis, and for identifying exons in a gene, particularly
CC using human lung derived mRNA and for the study of lung diseases
CC such as asthma, lung cancer, chronic obstructive pulmonary disease
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
CC pulmonary alveolar proteinosis, Karsagen syndrome, fibrocystic
CC pulmonary alveolar proteinosis, primary ciliary dyskinesia, pulmonary hypertension
CC and hyaline membrane disease. The present sequence is a single exon
CC probe open reading frame of the invention.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 422 BP; 53 A; 145 C; 159 G; 65 T; 0 other;

Query Match 55.6%; Score 422; DB 24; Length 422;

Best Local Similarity 100.0%; Pred. No. 3e-186;
Matches 422; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 ACCTGTCGAGTACAGTCCGCGAGTGCATTCACCCGCTACGTCGATGGCCGT 316
DB 422 ACCTGTCGAGTACAGTCCGCGAGTGCATTCACCCGCTACGTCGATGGCCGT 363
QY 317 GCCGACGCGCAGCGGATGACCGAGTGTGTCTCCGCGCAGTGTGACCCGCGCC 376
DB 362 GCCGACGCGCAGCGGATGACCGAGTGTGTCTCCGCGCAGTGTGACCCGCGCC 303
QY 377 TGTGTCGCAACGCGATGCGCGCGGCAAGTGTGTGCGACCTAGTGGCCCTCGACTTC 436
DB 302 TGTGTCGCAACGCGATGCGCGCGGCAAGTGTGTGCGACCTAGTGGCCCTCGACTTC 243
QY 437 GCATCCCGACCGCTACCGCGCGACGCGTGAAGTGTGTCTCCGCTGTGTGAAGCCG 496
DB 242 GCATCCCGACCGCTACCGCGCGACGCGTGAAGTGTGTCTCCGCTGTGTGAAGCCG 183
QY 497 CGCGCGCGCAGGATGTGCGCTGTGTGCTGTGTGCAAGTGTGCAAGCCGCTTCC 556
DB 182 CGCGCGCGCAGGATGTGCGCTGTGTGCTGTGTGCAAGTGTGCAAGCCGCTTCC 123
QY 557 ACAACGATGAGCTCAGAGACTTGGGACCGAGGCGCTCGGCGCGCAGAGAGGCCGGA 616

DB 122 ACAACGATGAGCTCAGAGACTTGGGACCGAGGCGCTCGGCGCGCAGAGAGGCCGGA 63

QY 617 ACCCGCGCGCGCGCGCGCGCGCGAGCGCCAAAGCCATCAGCGCGAGTGTGAAGACGCTACT 676

DB 62 ACCCGCGCGCGCGCGCGCGCGCGAGCGCCAAAGCCATCAGCGCGAGTGTGAAGACGCTACT 3

QY 677 AG 678

DB 2 AG 1

RESULT 15

AAA94050
ID AAA94050 standard; DNA; 1104 BP.

AAA94050;

15-JAN-2001 (first entry)

Human DAN/Cerberus-related protein 6 (hDCR6) coding sequence #1.

Human, DNA/Cerberus-related protein 6; hDCR6; morphogenic protein;
antagonist; BMP; cell growth; cell differentiation; bone formation;
gene therapy; ss.

Homo sapiens.

Key

Location/Qualifiers

CDS 1..1104

/tag= a

/product= "hDCR6 #1"

WO20005193-A2.

21-SEP-2000.

02-MAR-2000; 2000WO-US05337.

12-MAR-1999; 99US-0124118.

(REG-) REGENERON PHARM INC.

Economides AN;

WPI: 2000-638179/61.

P-PSDB; AAB26105.

Novel isolated, human DNA/Cerberus related protein 6 which include
natural homologue, and polypeptides comprising DCR6 domain and nucleic
acids encoding the proteins which are useful as probes and primers -

Claim 2; Fig 2; 40bp; English.

The present sequence comprises the human DAN/Cerberus-related protein 6
(hDCR6) coding sequence. It was isolated from a genomic DNA clone
following identification using computer-based 'virtual cloning'. hDCR6
is closely related to the DAN and DCR5 proteins, both of which act as
antagonists of morphogenic proteins such as BMP. It is possible that
the hDCR6 gene and protein can be used as immunogens, modulators of cell
function, growth and differentiation, to reduce undesirable bone
formation, to identify DCR6 binding agents, in diagnosis, and in gene
therapy.

Sequence 1104 BP; 226 A; 354 C; 364 G; 160 T; 0 other;

Query Match 55.6%; Score 422; DB 21; Length 1104;

Best Local Similarity 100.0%; Pred. No. 2.7e-186;

Matches 422; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 257 ACCTGTCGAGTACAGTCCGCGAGTGCATTCACCCGCTACGTCGATGGCCGT 316

DB 683 ACCTGTCGAGTACAGTCCGCGAGTGCATTCACCCGCTACGTCGATGGCCGT 742

Qy 317 GCCGACGCGCAAGCCGCTACCGAGCTGGTGTGCTCCGGCCAGTGGGCCCCGGCGGCC 376
|||
Db 743 GCCGACGCGCAAGCCGCTACCGAGCTGGTGTGCTCCGGCCAGTGGGCCCCGGCGGCC 802
|||
Qy 377 TGTGCTCCAAAGCCATCGGCGCGGCAAGTGTGGGCACTTGTGGGCCCGACTTCCGCT 436
|||
Db 803 TGTGCTCCAAAGCCATCGGCGCGGCAAGTGTGGGCACTTGTGGGCCCGACTTCCGCT 862
|||
Qy 437 GCATCCCGGACCGCTACCGGCGCGGCAAGTGTGGGCACTTGTGGGCCCGACTTCCGCT 496
|||
Db 863 GCATCCCGGACCGCTACCGGCGCGGCAAGTGTGGGCACTTGTGGGCCCGACTTCCGCT 922
|||
Qy 497 CGCGCGCGCGCAAGTGGCTGTGGCTGTGCAAGTGAAGCGCTCACCGCTTCC 556
|||
Db 923 CGCGCGCGCGCAAGTGGCTGTGGCTGTGCAAGTGAAGCGCTCACCGCTTCC 982
|||
Qy 557 ACAACCAAGTCGAGCTCAAGGACTTGGGACCGAGGCCGCTCGGCGCGAGAGGGCCGGA 616
|||
Db 983 ACAACCAAGTCGAGCTCAAGGACTTGGGACCGAGGCCGCTCGGCGCGAGAGGGCCGGA 1042
|||
Qy 617 AGCGCGCGCGCGCGCGCGCGCGCAAGCGCAAGCGAGCGCGAGTGGAGAGAGCGCTACT 676
|||
Db 1043 AGCGCGCGCGCGCGCGCGCGCGCAAGCGCAAGCGAGCGCGAGTGGAGAGAGCGCTACT 1102
|||
Qy 677 AG 678
||
Db 1103 AG 1104

Search completed: March 29, 2003, 08:39:17
Job time : 247.266 secs


```
Db 312 GTGACGATGGGCGCTGCGCGAGCGCAAGCCGAGTCAAGAGTGTGTCTCCGGCCAG 371
Qy 361 TGGGCGCCCGGCGCGCTGCTGCTGCCCAAGCCATGCGCGCGCAAGTGTGTGCACTAGT 420
Db 372 TGGGCGCCCGGCGCGCTGCTGCTGCCCAAGCCATGCGCGCGCAAGTGTGTGCACTAGT 431
Qy 421 GGGCGCCGACTTCCGCTGCTGCTGCCCAAGCCATGCGCGCGCAAGTGTGTGCACTAGT 480
Db 432 GGGCGCCGACTTCCGCTGCTGCTGCCCAAGCCATGCGCGCGCAAGTGTGTGCACTAGT 491
Qy 481 CCGGCTGCTGAGGCGCGCGCGCGCGCGCAAGTGTGTGCTGCTGCTGCTGCTGCTGCTGCT 540
Db 492 CCGGCTGCTGAGGCGCGCGCGCGCGCGCAAGTGTGTGCTGCTGCTGCTGCTGCTGCTGCT 551
Qy 541 CGGCTCAACCGCTTCCCAACCAAGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Db 552 CGGCTCAACCGCTTCCCAACCAAGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 611
Qy 601 CCGGAGAAAGGCGCGCGCGCGCGCGCGCGCGCAAGCCCAAGCCCAAGCCCAAGCCCAAG 660
Db 612 CCGGAGAAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 671
Qy 661 CTGGAGAACGCTACTAGAGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 720
Db 672 CTGGAGAACGCTACTAGAGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 731
Qy 721 GAACCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 759
Db 732 GAACCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 770
```

RESULT 2

```
US-09-449-218D-3
; Sequence 3, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 2301
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-449-218D-3
```

```
Query Match 93.3%; Score 708; DB 4; Length 2301;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 758; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 TACTGGAAGGTGGGCGTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
Db 12 TACTGGAAGGTGGGCGTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 71
Qy 61 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
Db 72 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 131
Qy 121 TTCAAGATGATGCGCAGGAATCATCCCGAGCTCGAGAGTACCCCGAGCTTCCACCG 180
Db 132 TTCAAGATGATGCGCAGGAATCATCCCGAGCTCGAGAGTACCCCGAGCTTCCACCG 191
```

```
Qy 181 GAGCTGGAACAACAAGACCATGAACCGGCGGAGAACGAGGGCGGCTCCCTCCACAC 240
Db 192 GAGCTGGAACAACAAGACCATGAACCGGCGGAGAACGAGGGCGGCTCCCTCCACAC 251
Qy 241 CCCTTTGAGACCAAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Db 252 CCCTTTGAGACCAAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 311
Qy 301 GTGACCAATGAGGCGGCTGCGCGAGCGGCGGAGCGGCTGCAAGGCTGTGTGCTGCTGCTGCTGCT 360
Db 312 GTGACCAATGAGGCGGCTGCGCGAGCGGCGGAGCGGCTGCAAGGCTGTGTGCTGCTGCTGCTGCT 371
Qy 361 TGGGCGCCCGGCGCGCTGCTGCTGCCCAAGCCATGCGCGCGCAAGTGTGTGCACTAGT 420
Db 372 TGGGCGCCCGGCGCGCTGCTGCTGCCCAAGCCATGCGCGCGCAAGTGTGTGCACTAGT 431
Qy 421 GGGCGCCGACTTCCGCTGCTGCTGCCCAAGCCATGCGCGCGCAAGTGTGTGCACTAGT 480
Db 432 GGGCGCCGACTTCCGCTGCTGCTGCCCAAGCCATGCGCGCGCAAGTGTGTGCACTAGT 491
Qy 481 CCGGCTGCTGAGGCGCGCGCGCGCGCGCGCGCGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Db 492 CCGGCTGCTGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 551
Qy 541 CGGCTCAACCGCTTCCCAACCAAGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Db 552 CGGCTCAACCGCTTCCCAACCAAGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 611
Qy 601 CCGGAGAAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
Db 612 CCGGAGAAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 671
Qy 661 CTGGAGAACGCTACTAGAGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 720
Db 672 CTGGAGAACGCTACTAGAGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 731
Qy 721 GAACCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 759
Db 732 GAACCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 770
```

RESULT 3

```
US-09-449-218D-7
; Sequence 7, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 2301
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-449-218D-7
```

```
Query Match 93.3%; Score 708; DB 4; Length 2301;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 758; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 TACTGGAAGGTGGGCGTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
Db 12 TACTGGAAGGTGGGCGTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 71
```


NUMBER OF SEQ ID NOS: 45
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 9
LENGTH: 642
TYPE: DNA
ORGANISM: Cercopithecus pygerythrus
US-09-449-218D-9

Query Match 14.5%; Score 110; DB 4; Length 642;
Best Local Similarity 99.1%; Pred. No. 1.5e-41;
Matches 210; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 121 TTCAAGATGATGCCAAGAAATCATCCCGAGCTCGAGAGTACCCGAGCTCCACCG 180
DB 85 TTCAAGATGATGCCAAGAAATCATCCCGAGCTCGAGAGTACCCGAGCTCCACCG 144
QY 181 GAGCTGGAGAACAAAGACATGAAACCGGCGGAGAACGAGAGGGGGCTTCCCAACAC 240
DB 145 GAGCTGGAGAACAAAGACATGAAACCGGCGGAGAACGAGAGGGGGCTTCCCAACAC 204
QY 241 CCCTTTGAGACCAAAAGACGTGTCGAGTACAGCTGCGGAGCTGACATTACCGGCTAC 300
DB 205 CCCTTTGAGACCAAAAGACGTGTCGAGTACAGCTGCGGAGCTGACATTACCGGCTAC 264
QY 301 GTGACCGATGGCGCGTGCAGCGGACGCGCAAGCC 332
DB 265 GTGACCGATGGCGCGTGCAGCGGACGCGCAAGCC 296

RESULT 8
US-09-449-218D-15
Sequence 15, Application US/09449218D
Patent No. 6395511

GENERAL INFORMATION:
APPLICANT: Brunkow, Mary E.
APPLICANT: Galas, David J.
APPLICANT: Kovacevich, Brian
APPLICANT: Mulligan, John T.
APPLICANT: Paepert, Bryan W.
APPLICANT: Van Ness, Jeffrey
APPLICANT: Winkler, David G.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
FILE REFERENCE: 240083.508
CURRENT APPLICATION NUMBER: US/09/449,218D
CURRENT FILING DATE: 1999-11-24
NUMBER OF SEQ ID NOS: 45
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 15
LENGTH: 532
TYPE: DNA
ORGANISM: Bos taurus
US-09-449-218D-15

Query Match 10.9%; Score 83; DB 4; Length 532;
Best Local Similarity 99.3%; Pred. No. 3.7e-29;
Matches 133; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 280 GAGCTGACATTCACCCCGCTACGATGAGCGGCTGCGGACGCCCAAGCGGTACAC 339
DB 153 GAGCTGACATTCACCCCGCTACGATGAGCGGCTGCGGACGCCCAAGCGGTACAC 212
QY 340 GAGCTGATGCTCGGAGCAGTGGCGCGGCGGCTGCTGCCCAAGACCATCGAGCGCG 399
DB 213 GAGCTGATGCTCGGAGCAGTGGCGCGGCGGCTGCTGCCCAAGACCATCGAGCGCG 272
QY 400 GGCAGGTGTGGCG 413
DB 273 GGCAGGTGTGGCG 286

RESULT 9
US-09-449-218D-11

Sequence 11, Application US/09449218D
Patent No. 6395511
GENERAL INFORMATION:
APPLICANT: Brunkow, Mary E.
APPLICANT: Galas, David J.
APPLICANT: Kovacevich, Brian
APPLICANT: Mulligan, John T.
APPLICANT: Paepert, Bryan W.
APPLICANT: Van Ness, Jeffrey
APPLICANT: Winkler, David G.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
FILE REFERENCE: 240083.508
CURRENT APPLICATION NUMBER: US/09/449,218D
CURRENT FILING DATE: 1999-11-24
NUMBER OF SEQ ID NOS: 45
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 11
LENGTH: 638
TYPE: DNA
ORGANISM: Mus musculus
US-09-449-218D-11

Query Match 9.1%; Score 69; DB 4; Length 638;
Best Local Similarity 100.0%; Pred. No. 9.7e-23;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 517 CTGTGGGCTGTGTGAATGCAAGGCGCTTCCACCAACCACTCGAGTCAAG 576
DB 475 CTGTGGGCTGTGTGAATGCAAGGCGCTTCCACCAACCACTCGAGTCAAG 534
QY 577 GACTTCGGG 585
DB 535 GACTTCGGG 543

RESULT 10
US-09-449-218D-17
Sequence 17, Application US/09449218D
Patent No. 6395511
GENERAL INFORMATION:
APPLICANT: Brunkow, Mary E.
APPLICANT: Galas, David J.
APPLICANT: Kovacevich, Brian
APPLICANT: Mulligan, John T.
APPLICANT: Paepert, Bryan W.
APPLICANT: Van Ness, Jeffrey
APPLICANT: Winkler, David G.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
FILE REFERENCE: 240083.508
CURRENT APPLICATION NUMBER: US/09/449,218D
CURRENT FILING DATE: 1999-11-24
NUMBER OF SEQ ID NOS: 45
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 17
LENGTH: 35828
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(35828)
OTHER INFORMATION: n = A,T,C or G
US-09-449-218D-17

Query Match 9.1%; Score 69; DB 4; Length 35828;
Best Local Similarity 100.0%; Pred. No. 7.8e-23;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 517 CTGTGGGCTGTGTGAATGCAAGGCGCTTCCACCAACCACTCGAGTCAAG 576
DB 21283 CTGTGGGCTGTGTGAATGCAAGGCGCTTCCACCAACCACTCGAGTCAAG 21342

577 GACTTCGGG 585
 |||||
 21343 GACTTCGGG 21351
 Db

RESULT 11
 US-09-449-218D-13
 : Sequence 13 Application US/09449218D
 : Patent No. 6395511
 : GENERAL INFORMATION:
 : APPLICANT: Brunkow, Mary E.
 : APPLICANT: Galas, David J.
 : APPLICANT: Kovacevich, Brian
 : APPLICANT: Mulligan, John T.
 : APPLICANT: Paepert, Bryan W.
 : APPLICANT: Van Ness, Jeffrey
 : APPLICANT: Winkler, David G.
 : TITLE OF INVENTION: COMPOSITIONS AND METHODS
 : OF TITLE OF INVENTION: BONE MINERALIZATION
 : FILE REFERENCE: 240083.508
 : CURRENT APPLICATION NUMBER: US/09/449,218D
 : CURRENT FILING DATE: 1999-11-24
 : NUMBER OF SEQ ID NOS: 45
 : SOFTWARE: FASTSEQ for Windows Version 3.0
 : SEQ ID NO 13
 : LENGTH: 674
 : TYPE: DNA
 : ORGANISM: Rattus norvegicus
 : US-09-449-218D-13

Query Match	9.0%;	Score 68;	DB 4;	Length 674;
Best Local Similarity	100.0%;	Pred. No. 2.8e-22;		
Matches 68;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

Qy	517	CTGTGGCCCTCGTGCAGGTGCAGGCGGCTTCACCGGTTCCACAACGAGTGGAGCTCAAG	516
Db	513	CTGTGGCCCTCGTGCAGGTGCAGGCGGCTTCACCGGTTCCACAACGAGTGGAGCTCAAG	517

Qy	577	GACTTCGG	584
Db	573	GACTTCGG	580

```

RESULT 12
US-09-449-218D-26
; Sequence 26, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepert, Bryan W.
; APPLICANT: Van Nees, Jeffrey
; APPLICANT: Whinkle, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHOD
; TITLE OF INVENTION: BONE MINERALIZATION
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218BD
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 39
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer for PCR
US-09-449-218D-26

```

Query Match	3.3%;	Score 25;	DB 4;	Length 39;
Best Local Similarity	100.0%;	Pred. No. 0.018;		
Matches	25;	Conservative	0;	Mismatches 0;
			Indels	0;
			Gaps	0

Qy 108 GGGGTGGCAGGCGTTCAAGATGAT 132
 Db 15 GGGGTGGCAGGCGTTCAAGATGAT 39

```

RESULT 13
US-09-449-218D-19
: Sequence 19, Application US/09449218D
: Patent No. 6395511
: GENERAL INFORMATION:
: APPLICANT: Brunkow, Mary E.
: APPLICANT: Galas, David J.
: APPLICANT: Kovacevich, Brian
: APPLICANT: Mulligan, John T.
: APPLICANT: Paepert, Bryan W.
: APPLICANT: Van Ness, Jeffrey
: APPLICANT: Winkler, David G
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
: TITLE OF INVENTION: BONE MINERALIZATION
: FILE REFERENCE: 240083.508
: CURRENT APPLICATION NUMBER: US/09/449,218D
: CURRENT FILING DATE: 1999-11-24
: NUMBER OF SEQ ID NOS: 45
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 19
: LENGTH: 21
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Primer for PCR
US-09-449-218D-19

```

Query Match	2.8%;	Score 21;	DB 4;	Length 21;
Best Local Similarity	100.0%;	Pred. No. 1.3;		
Matches 21;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 178 CCGAGCTGGAGACAACAAG 198
|||
Db 1 CCGAGCTGGAGACAACAAG 21

```

RESULT 14
US-09-449-218D-27/C
Sequence 27, Application US/09449218D
Patent No. 6395511
GENERAL INFORMATION:
APPLICANT: Brunkow, Mary E.
APPLICANT: Galas, David J.
APPLICANT: Kovacevich, Brian
APPLICANT: Mulligan, John T.
APPLICANT: Paepert, Bryan W.
APPLICANT: Van Ness, Jeffrey
APPLICANT: Winkler, David G.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
FILE REFERENCE: 240083.508
CURRENT APPLICATION NUMBER: US/09/449,218D
CURRENT FILING DATE: 1999-11-24
NUMBER OF SEQ ID NOS: 45
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 27
LENGTH: 57
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE: Primer for PCR
US-09-449-218D-27

```

```
Query Match      2.8%; Score 21; DB 4; Length 57;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0
```

QY 655 GCCGAGCTGGAGAACGCTTAC 675
 |||||
 Db 57 GCCGAGCTGGAGAACGCTTAC 37

RESULT 15

US-09-213-767-1
 ; Sequence 1, Application US/09213767
 ; Patent No. 5948680
 ; GENERAL INFORMATION:
 ; APPLICANT: Brenda F. Baker
 ; APPLICANT: Lex M. Cowsett
 ; TITLE OF INVENTION: ANTISENSE MODULATION OF ELK-1 EXPRESSION
 ; FILE REFERENCE: RTS-0024
 ; CURRENT APPLICATION NUMBER: US/09/213,767
 ; CURRENT FILING DATE: 1998-12-17
 ; NUMBER OF SEQ ID NOS: 47
 ; SEQ ID NO 1
 ; LENGTH: 2266
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (316)..(1602)
 US-09-213-767-1

Query Match 2.8%; Score 21; DB 2; Length 2266;
 Best Local Similarity 100.0%; Pred. No. 0.98;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 600 GCCGCAAGAGGGCCGGAGGCC 620
 |||||
 Db 1242 GCCGCAAGAGGGCCGGAGGCC 1262

Search completed: March 29, 2003, 10:46:53
 JOD time : 66.7914 secs




```

QY 241 CCCTTGAGACAAAGAGCTGTCCGAGTACAGTCCGCGAGCTGACCTTACCCGCTAC 300
D 241 CCCTTGAGACAAAGAGCTGTCCGAGTACAGTCCGCGAGCTGACCTTACCCGCTAC 300
QY 301 GTGACCGATGGGCGGTGCGCGAGCGCAAGCGCGGTGCAACGAGTGTGTCTCCGGCAG 360
D 301 GTGACCGATGGGCGGTGCGCGAGCGCAAGCGCGGTGCAACGAGTGTGTCTCCGGCAG 360
QY 361 TGGCGCCCGGCGCGCTGTGTCTGTCCCAAGCCATCGCGCGCGCAAGTGTGTGCGAAGT 420
D 361 TGGCGCCCGGCGCGCTGTGTCTGTCCCAAGCCATCGCGCGCGCAAGTGTGTGCGAAGT 420
QY 421 GGGCGCCGACTTCCGCTGATCCCGACCGCTACCGCGCGAGCGCGTGTGTGTGT 480
D 421 GGGCGCCGACTTCCGCTGATCCCGACCGCTACCGCGCGAGCGCGTGTGTGTGTGT 480
QY 481 CCCGCTGTGAGGCGCGCGCGCGCGCAAGGTGCGCTGTGTGTGTGTGTGTGTGTGT 540
D 481 CCCGCTGTGAGGCGCGCGCGCGCGCAAGGTGCGCTGTGTGTGTGTGTGTGTGTGT 540
QY 541 CGCCTCACCCGCTTCCACACAGTCCGAGCTCAAGACTTCCGAGACCGAGCGCGCTCGG 600
D 541 CGCCTCACCCGCTTCCACACAGTCCGAGCTCAAGACTTCCGAGACCGAGCGCGCTCGG 600
QY 601 CCGGAGAAAGGCGCGAAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
D 601 CCGGAGAAAGGCGCGAAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
QY 661 CTGAGAAAGCGCTTACAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 720
D 661 CTGAGAAAGCGCTTACAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 720
QY 721 GAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 759
D 721 GAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 759

```

RESULT 2

```

US-09-864-761-30988/c
; Sequence 30988, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668

```

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 30988
; LENGTH: 422
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC003098.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.62
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.49
; OTHER INFORMATION: BEST HUMAN HIT: BE613498.1, EVALUATE 9.90e-02
; OTHER INFORMATION: SWISSPROT HIT: P45646, EVALUATE 4.90e-01
; OTHER INFORMATION: NT HIT: Af074705.1, EVALUATE 1.00e+00

```

Query Match

```

Best Local Similarity 100.0%; Pred. No. 6e-214;
Matches 422; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 257 ACGTGTCCGAGTACAGCTGCTCCGCGAGCTGCACTTCAACCGCTACGTAAGCGGCGCT 316
D 422 ACGTGTCCGAGTACAGCTGCTCCGCGAGCTGCACTTCAACCGCTACGTAAGCGGCGCT 363
QY 317 GCCGAGGCGCCAAAGCGGCTGCAACGAGTGTGTGTCTCCGCGCAATGCGCGCGCGCGCC 376
D 362 GCCGAGGCGCCAAAGCGGCTGCAACGAGTGTGTGTCTCCGCGCAATGCGCGCGCGCGCC 303
QY 377 TCGTGTCCCAAGCGCATGCGCGCGCGCGCAAGTGTGTGTGTCTCCGCGCAATGCGCGCGCGCT 436
D 302 TCGTGTCCCAAGCGCATGCGCGCGCGCGCAAGTGTGTGTGTCTCCGCGCAATGCGCGCGCGCT 243
QY 437 GCATCCCGGACCGCTACCGCGCGAGCGCGGTGCAAGCTGTGTGTGTGTGTGTGTGTGTGT 496
D 242 GCATCCCGGACCGCTACCGCGCGAGCGCGGTGCAAGCTGTGTGTGTGTGTGTGTGTGTGT 183
QY 497 CCGCGCGCGCGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 556
D 182 CCGCGCGCGCGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 123
QY 557 ACAACCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 616
D 122 ACAACCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 63
QY 617 AGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 676
D 62 AGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 3
QY 677 AG 678
D 2 AG 1

```

RESULT 3

```

US-09-864-761-14440/c
; Sequence 14440, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:

```

```

APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DETERMINED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aemita-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 14440
LENGTH: 392
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC003098.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.62
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.6
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.49
US-09-864-761-14440

Query Match          46.8%; Score 355; DB 10; Length 392;
Best Local Similarity 100.0%; Pred. No. 2e-178;
Matches 355; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

585 GACCGAGCCGCTCGGCGCAGAAAGGCGGAGCCGCGCCCGCGGAGCGCAA 644
212 GACCGAGCCGCTCGGCGCAGAAAGGCGGAGCCGCGCCCGCGGAGCGCAA 153
645 AGCCACCAAGCCGAGCTGAGAAAGCCTTACTAGAGCCGCGCGCCCTTCCACCGG 704
152 AGCCACCAAGCCGAGCTGAGAAAGCCTTACTAGAGCCGCGCGCCCTTCCACCGG 93
705 CGGCGCGCCCGCGCCTGAAACCGGCGCCCACTTCTGCTCGCGGCTTT 759
92 CGGCGCGCCCGCGCCTGAAACCGGCGCCCACTTCTGCTCGCGGCTTT 38

RESULT 4
Sequence 3, Application US/09867274
Patent No. US20020106650A1
GENERAL INFORMATION:
APPLICANT: Paszty, Christopher
TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
FILE REFERENCE: 01017/37428
CURRENT APPLICATION NUMBER: US/09/867,274
PRIOR APPLICATION NUMBER: US 60/208,550
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: US 60/223,542
PRIOR FILING DATE: 2000-08-04
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3
LENGTH: 636
TYPE: DNA
ORGANISM: Mus musculus
US-09-867-274-3

Query Match          9.1%; Score 69; DB 10; Length 636;
Best Local Similarity 100.0%; Pred. No. 7.4e-27;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

517 CTGTGGCTCTGTGCAAGTGAAGCGGCTTCCAGCAACGATGCGAGCTCAAG 576
475 CTGTGGCTCTGTGCAAGTGAAGCGGCTTCCAGCAACGATGCGAGCTCAAG 534
577 GACTTCGGG 585
535 GACTTCGGG 543

RESULT 5
US-09-867-274-21
Sequence 21, Application US/09867274
Patent No. US20020106650A1
GENERAL INFORMATION:
APPLICANT: Paszty, Christopher
TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
FILE REFERENCE: 01017/37428
CURRENT APPLICATION NUMBER: US/09/867,274
PRIOR APPLICATION NUMBER: US 60/208,550
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: US 60/223,542
PRIOR FILING DATE: 2000-08-04
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn version 3.0
SEQ ID NO 21
LENGTH: 27
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Artificial; PCR primer

```

US-09-867-274-21

Query Match 3.6%; Score 27; DB 10; Length 27;
Best Local Similarity 100.0%; Pred.No.0.00019;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 TGTGTCGTCTGCGCTGCTGTACACA 82
DB 1 TGTGTCGTCTGCGCTGCTGTACACA 27

RESULT 6

US-09-867-274-8/c
; Sequence 8, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; APPLICANT: Gao, Yongming
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 26
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial: PCR primer
US-09-867-274-8

Query Match 3.4%; Score 26; DB 10; Length 26;
Best Local Similarity 100.0%; Pred.No.0.00064;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 734 ACATTTCGTCTGCGCGCTGTTT 759
DB 26 ACATTTCGTCTGCGCGCTGTTT 1

RESULT 7

US-09-867-274-7
; Sequence 7, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; APPLICANT: Gao, Yongming
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 24
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial: PCR primer
US-09-867-274-7

Query Match 3.2%; Score 24; DB 10; Length 24;
Best Local Similarity 100.0%; Pred.No.0.0074;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TACTGGAAGTGGCGTCCCTCT 24
DB 1 TACTGGAAGTGGCGTCCCTCT 24

RESULT 8

US-09-867-274-22/c
; Sequence 22, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; APPLICANT: Gao, Yongming
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 23
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial: PCR primer
US-09-867-274-22

Query Match 3.0%; Score 23; DB 10; Length 23;
Best Local Similarity 100.0%; Pred.No.0.025;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 410 GCGACCTAGTGGCCGACTTC 432
DB 23 GCGACCTAGTGGCCGACTTC 1

RESULT 9

US-09-867-274-20/c
; Sequence 20, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; APPLICANT: Gao, Yongming
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; CURRENT FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 20
; LENGTH: 41
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial: PCR primer
US-09-867-274-20

Query Match 2.8%; Score 21; DB 10; Length 41;
Best Local Similarity 100.0%; Pred.No.0.27;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 658 GAGCTGAGAGCGCTACTAG 678
DB 36 GAGCTGAGAGCGCTACTAG 16


```
RESULT 10
US-09-815-242-5978/c
; Sequence 5978, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5978
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(645)
US-09-815-242-5978

Query Match          2.6%; Score 20; DB 10; Length 645;
Best Local Similarity 100.0%; Pred. No. 0.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 625 CCCCCGGGGGGAGGCCCA 644
DB 36 CCCCCGGGGGGAGGCCCA 17

RESULT 11
US-09-815-242-7536/c
; Sequence 7536, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9619
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Salmonella typhi
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(645)
```

```

; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7536
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(645)
US-09-815-242-7536

Query Match          2.5%; Score 19; DB 10; Length 645;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 626 CCCCCGGGGAGGCCCA 644
DB 35 CCCCCGGGGAGGCCCA 17

RESULT 12
US-09-815-242-9619/c
; Sequence 9619, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9619
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Salmonella typhi
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(645)
```

US-09-815-242-9619

Query Match 2.5%; Score 19; DB 10; Length 645;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 626 CCCGCGCCGAGCGCCAA 644
DB 35 CCCGCGCCGAGCGCCAA 17

RESULT 13
US-09-903-376-1
; Sequence 1, Application US/09903376
; Publication No. US20030023998A1
; GENERAL INFORMATION:
; APPLICANT: Brennan, Thomas J.
; TITLE OF INVENTION: TRANSGENIC MICE CONTAINING 5-HT-2B GENE
; FILE REFERENCE: R-599
; CURRENT APPLICATION NUMBER: US/09/903,376
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/218,358
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/223,120
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 60/223,122
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1550
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-903-376-1

Query Match 2.5%; Score 19; DB 9; Length 1550;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 CTCCCACTGGCCTGTGTC 61
DB 399 CTCCCACTGGCCTGTGTC 417

RESULT 14
US-09-867-274-10/c
; Sequence 10, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Paszty, Christopher
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 0101/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 24
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial: PCR primer
US-09-867-274-10

Query Match 2.4%; Score 18; DB 10; Length 24;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 426 CGACTTCGGCTGATCCC 443
DB 24 CGACTTCGGCTGATCCC 7

RESULT 15
US-09-864-761-3211/c
; Sequence 3211, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aesomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 3211
; LENGTH: 375
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL121752.2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 22
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 8.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.7
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.8

; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 7.2
US-09-864-761-3211

Query Match 2.4%; Score 18; DB 10; Length 375;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 464 GCCTGCAGCTGCTGTGTC 481
|||
Db 213 GCCTGCAGCTGCTGTGTC 196

Search completed: March 29, 2003, 10:49:31
Job time : 94.4064 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2003, 01:29:00 ; Search time 1634.43 Seconds
(without alignments)
7520.874 Million cell updates/sec

Title: US-09-867-274-1

Perfect score: 759
Sequence: 1 tactggaagtcgctgccc.....ctgtctctgcgcgtgttt 759

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST: *
1: em_estba: *
2: em_esthum: *
3: em_estin: *
4: em_estmu: *
5: em_estrov: *
6: em_estro: *
7: em_estro: *
8: em_hic: *
9: gb_est1: *
10: gb_est2: *
11: gb_hic: *
12: gb_est3: *
13: gb_est4: *
14: gb_est5: *
15: em_estfun: *
16: em_estom: *
17: gb_gss: *
18: em_gss_hum: *
19: em_gss_inv: *
20: em_gss_pln: *
21: em_gss_vit: *
22: em_gss_fun: *
23: em_gss_mam: *
24: em_gss_mus: *
25: em_gss_other: *
26: em_gss_pro: *
27: em_gss_yod: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	232	30.6	535	17	AQ171546 HS_3088_B
2	69	9.1	623	10	BB637315
3	69	9.1	1990	11	AK017295 Mus muscu
4	68	9.0	291	10	BE111224
5	68	9.0	360	9	A1556282
6	68	9.0	419	10	BE101082

Result No.	Score	Query Match	Length	DB ID	Description
7	38	5.0	315	9	A1113131
8	31	4.1	511	10	BB638050
9	31	4.1	627	17	AZ876400
10	29	3.8	107	10	BB636457
11	26	3.4	147	17	AZ746508
12	25	3.3	368	12	BC711211
13	25	3.3	385	10	AM434705
14	21	2.8	472	12	BF420024
15	21	2.8	477	17	AQ0803172
16	21	2.8	489	13	AA782033
17	21	2.8	532	12	BF407514
18	21	2.8	566	13	BF461146
19	21	2.8	639	9	AL525954
20	21	2.8	659	13	BI755351
21	21	2.8	697	10	BE613498
22	21	2.8	698	13	BI459804
23	21	2.8	743	13	BI757671
24	21	2.8	758	13	BI160501
25	21	2.8	775	13	BM012875
26	21	2.8	806	12	BG818868
27	21	2.8	907	14	BQ737053
28	21	2.8	1026	13	BM475656
29	21	2.8	1049	14	BM905809
30	21	2.8	1115	14	BM805183
31	21	2.8	1168	14	BM802647
32	20	2.6	413	9	AA869749
33	20	2.6	420	9	AA981192
34	20	2.6	482	9	A1592910
35	20	2.6	516	9	AA794777
36	20	2.6	687	13	BI408475
37	20	2.6	719	14	BM951395
38	20	2.6	737	17	BI9523
39	20	2.6	739	17	AZ194950
40	20	2.6	740	12	BF204444
41	20	2.6	780	13	BI411733
42	20	2.6	990	11	AK002527
43	20	2.6	1062	17	CNS05C1W
44	20	2.6	1270	14	BQ218112
45	20	2.6	1282	12	BG482101

ALIGNMENTS

RESULT 1
AQ171546/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
COMMENT

535 bp DNA linear GSS 17-OCT-1998
HS_3088_B1_H06_MF CIT Approved Human Genomic Sperm Library D Homo
sapiens genomic clone Plate=3088 Col=11 Row=P, DNA sequence.
AQ171546
AQ171546.1 GI:3568913
GSS.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and
Hood, L.
Sequence-tagged connectors: A sequence approach to mapping and
examining the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Sequence Tagged Connector
Plate: 3088 Row: P Column: 11

ORGANISM Rattus norvegicus
 Bakayota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (bases 1 to 291)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 COMMENT Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.uiowa.edu
 Oligo-dt track not found, Not 1 site shown in beginning of sequence
 is likely internal to the message. cDNA library Preparation: M.B.
 Soares Lab Clone distribution: clones will be available through
 Research Genetics (www.resgen.com) The following repetitive
 elements were found in this cDNA sequence: 97-125,
 >GC_richflow complexity
 Seq primer: M13 Forward
 POLYA=No.

FEATURES
 source Location/Qualifiers
 1..291
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-BJ1-auw-b-04-0-UI"
 /clone_1lb="UI-R-BJ1"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pRT3D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-BJ1
 library is a subtracted library derived from the following
 tissues: atrium at 16.5 dpc, ventricle at 16.5 dpc, AV
 canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc,
 AV canal at 15 dpc, ventricle at 13 dpc, and adult heart.
 For a detailed description of the library from which this
 clone was derived, please visit our web site at
 ratseq.eng.uiowa.edu. The subtraction has been previously
 described in (Bonaldo, Lennon and Soares, Genome Research
 6:791-806, 1996)
 TAG_SEQ=None found"
 BASE COUNT 41 a 101 c 109 g 40 t

ORIGIN
 Query Match 9.0%; Score 68; DB 10; Length 291;
 Best Local Similarity 100.0%; Pred. No. 3.3e-22;
 Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 517 CTGTGGGCTGTGCAAGTGAAGCGGCTTCCACCAACGAGCTCAAG 576
 |||||
 DB 87 CTGTGGGCTGTGCAAGTGAAGCGGCTTCCACCAACGAGCTCAAG 28

QY 577 GACTTCGG 584
 |||||
 DB 27 GACTTCGG 20

RESULT 5
 A1556282/c 360 bp mRNA linear EST 23-MAR-1999
 LOCUS A1556282
 DEFINITION UI-R-C2p-rk-a-03-0-UI.s1 UI-R-C2p Rattus norvegicus cDNA clone
 accession A1556282
 version A1556282.1 GI:4488645
 keywords EST.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Bakayota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE 1 (bases 1 to 360)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery

JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 COMMENT Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.uiowa.edu
 Oligo-dt track not found, Not 1 site shown in beginning of sequence
 is likely internal to the message. cDNA library Preparation: M.
 Fatima Bonaldo, Ph.D. Clone distribution: clones will be available
 through Research Genetics (www.resgen.com) The following repetitive
 elements were found in this cDNA sequence: 97-125,
 >GC_richflow complexity
 Seq primer: M13 Forward

FEATURES
 source Location/Qualifiers
 1..360
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-C2p-rk-a-03-0-UI"
 /clone_1lb="UI-R-C2p"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pRT3D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-C2p
 library is a subtracted library derived from the UI-R-C1
 library, which is a subtracted library derived from the
 UI-R-C0 library. The UI-R-C0 library consisted of a
 mixture of individually tagged normalized libraries
 constructed from rat placenta, adult lung, brain, liver,
 kidney, heart, spleen, ovary, muscle, 8, 12 and 18-day
 embryo. The tag is a string of 3-5 nucleotides present
 between the Not I site and the oligo-dt track which allows
 identification of the library of origin of a clone within
 the mixture. The subtracted library (UI-R-C2p) was
 constructed as follows: PCR amplified cDNA inserts from
 UI-R-C1 clones from which 3' ESTs had been derived was
 used as a driver in a hybridization with the UI-R-C1
 library in the form of single-stranded circles. The
 remaining single-stranded circles (subtracted library) was
 purified by hydroxyapatite column chromatography,
 converted to double-stranded circles and electroporated
 into DH10B bacteria (Life Technologies) to generate the
 UI-R-C2p library. This procedure has been previously
 described (Bonaldo, Lennon and Soares, Genome Research 6:
 791-806, 1996)"

BASE COUNT 54 a 120 c 131 g 54 t 1 others

ORIGIN
 Query Match 9.0%; Score 68; DB 9; Length 360;
 Best Local Similarity 100.0%; Pred. No. 3.5e-22;
 Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 517 CTGTGGGCTGTGCAAGTGAAGCGGCTTCCACCAACGAGCTCAAG 576
 |||||
 DB 87 CTGTGGGCTGTGCAAGTGAAGCGGCTTCCACCAACGAGCTCAAG 28

QY 577 GACTTCGG 584
 |||||
 DB 27 GACTTCGG 20

RESULT 6
 BE101082/c 419 bp mRNA linear EST 13-JUN-2000
 LOCUS BE101082
 DEFINITION UI-R-BJ1-ety-h-10-0-UI.s1 UI-R-BJ1 Rattus norvegicus cDNA clone

ACCESSION UI-R-BJ1-aty-h-10-0-UI 3', mRNA sequence.
 VERSION BE101082
 SOURCE BE101082.1 GI:84922989
 KEYWORDS EST.
 ORGANISM Norway rat.
 SOURCE Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 419)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 COMMENT Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.uiowa.edu
 Oligo-dT track not found, Not 1 site shown in beginning of sequence
 is likely internal to the message. cDNA Library Preparation: M.B.
 Soares Lab Clone distribution: clones will be available through
 Research Genetics (www.resgen.com) The following repetitive
 elements were found in this cDNA sequence: 97-125,
 >GC_rich#low_complexity
 Seq primer: M13 Forward
 POLYA=No.

FEATURES
 source
 1..419 Location/Qualifiers
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-BJ1-aty-h-10-0-UI"
 /clone_11b="UI-R-BJ1"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pRT73D-Pac (Pharmacia) with a modified
 polylinker. Site_1: Not 1; Site_2: Eco RI; The UI-R-BJ1
 library is a subtracted library derived from the following
 tissues: atrium at 16.5 dpc, ventricle at 16.5 dpc, AV
 canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc,
 AV canal at 15 dpc, ventricle at 13 dpc, and adult heart.
 For a detailed description of the library from which this
 clone was derived, please visit our web site at
 rategen.uiowa.edu. The subtraction has been previously
 described in (Bonaldo, Lennon and Soares, Genome Research
 6:791-806, 1996)
 TAG_SEQ=None found"
 BASE COUNT 56 a 135 c 155 g 73 t
 ORIGIN

Query Match 9.0%; Score 68; DB 10; Length 419;
 Best Local Similarity 100.0%; Pred. No. 3.5e-22;
 Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 517 CTGTGGCTCTGTGCAAGTGCAGCGCTTCACACCAAGTGCAGCTCAAG 576
 |||||||
 DB 87 CTGTGGCTCTGTGCAAGTGCAGCGCTTCACACCAAGTGCAGCTCAAG 28
 |||||||
 QY 577 GACTTCGG 584
 |||||||
 DB 27 GACTTCGG 20

RESULT 7
 A1113131/c 315 bp mRNA linear EST 11-FEB-1999
 LOCUS UI-R-C2p-nr-f-06-0-UI.61 UI-R-C2p Rattus norvegicus cDNA clone
 DEFINITION UI-R-C2p-nr-f-06-0-UI 3', mRNA sequence.
 ACCESSION A1113131
 VERSION A1113131.1 GI:3513080

KEYWORDS EST.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE 1 (bases 1 to 315)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 COMMENT Contact: Soares, MB
 Program for Rat Gene Discovery and Mapping
 University of Iowa
 451 Eckstein Medical Research Building Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: msoares@blue.weeg.uiowa.edu
 Oligo-dT track not found, Not 1 site shown in beginning of sequence
 is likely internal to the message. cDNA Library Preparation: M.
 Fatima Bonaldo, Ph.D. Clone distribution: clones will be available
 through Research Genetics The following repetitive elements were
 found in this cDNA sequence: 97-125, >GC_rich#low_complexity
 Seq primer: M13 Forward.

FEATURES
 source
 1..315 Location/Qualifiers
 /organism="Rattus norvegicus"
 /strain="Sprague-Dawley"
 /db_xref="taxon:10116"
 /clone="UI-R-C2p-nr-f-06-0-UI"
 /clone_11b="UI-R-C2p"
 /dev_stage="adult"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pRT73D-Pac (Pharmacia) with a modified
 polylinker. Site_1: Not 1; Site_2: Eco RI; The UI-R-C2p
 library is a subtracted library derived from the UI-R-C1
 library, which is a subtracted library derived from the
 UI-R-C0 library. The UI-R-C0 library consisted of a
 mixture of individually tagged normalized libraries
 constructed from rat placenta, adult lung, brain, liver,
 kidney, heart, spleen, ovary, muscle, 8, 12 and 18-day
 embryo. The tag is a string of 3-5 nucleotides present
 between the Not 1 site and the oligo-dT track which allows
 identification of the library of origin of a clone within
 the mixture. The subtracted library (UI-R-C2p) was
 constructed as follows: PCR amplified cDNA inserts from
 UI-R-C1 clones from which 3' ESTs had been derived was
 used as a driver in a hybridization with the UI-R-C1
 library in the form of single-stranded circles. The
 remaining single-stranded circles (subtracted library) was
 purified by hydroxyapatite column chromatography,
 converted to double-stranded circles and electroporated
 into DH10B bacteria (Life Technologies) to generate the
 UI-R-C2p library. This procedure has been previously
 described (Bonaldo, Lennon and Soares, Genome Research 6:
 791-806, 1996)"

BASE COUNT 46 a 110 c 116 g 43 t
 ORIGIN

Query Match 5.0%; Score 38; DB 9; Length 315;
 Best Local Similarity 100.0%; Pred. No. 1.3e-07;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 518 TGTGTGGCTCTGTGCAAGTGCAGCGCTTCACCGCTTC 555
 |||||||
 DB 86 TGTGTGGCTCTGTGCAAGTGCAGCGCTTCACCGCTTC 49
 |||||||

RESULT 8
 BB638050 511 bp mRNA linear EST 31-AUG-2001
 LOCUS BB638050
 DEFINITION BB638050 RIKEN full-length enriched, 3 days neonate thymus Mus

ACCESSION	musculus cDNA A630006E24 5', mRNA sequence.
VERSION	B8638050
KEYWORDS	B8638050.1 GI:15401086
SOURCE	EST.
ORGANISM	house mouse. Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 511)
AUTHORS	Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Himoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda , M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sassaki , D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
TITLE	RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
JOURNAL	Unpublished (2001)
COMMENT	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsc.riken.go.jp/ URL:http://genome.gsc.riken.go.jp/ Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh , M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) wagui,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsunura , S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Konno,H., Fukuishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara, , Y. and Hayashizaki,Y. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) Yamanaka,I., Kitayosa,H., Kondo,S., Saito,T., Shinagawa,A., Atzawa, , K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K., Arakawa,T., Ishii,Y. and Hayashizaki,Y. Mapping of 1903 mouse cDNAs on mouse chromosomes. J. Struct. Func. Genomics 2 pre, L72-L86 (2001) Please visit our web site (http://genome.gsc.riken.go.jp) for further details. e mouse tissues. Location/Qualifiers 1..511 /organism="Mus musculus" /db_xref="taxon:10090" /clone_id="A630006E24" /clone_lib="RIKEN full-length enriched, 3 days neonate thymus" /tissue_type="Thymus" /dev_stage="3 days neonate" /lab_host="DH10B" /note=Site_1: SalI, Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGGAGAGAGATCCAAAGACTCTTTTCTTTTTTTTNN 3'], cDNA was prepared by using trihalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization

BASE COUNT	86 a	182 c	162 g	81 t
ORIGIN	to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence 15' GAGAGAGAGATTCGCGATTAAATTAATTAATCCCCCCCCCCCC 3'. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda PLC 1."			
Query Match	4.1%; Score 31; DB 10; Length 511;			
Best Local Similarity	100.0%; Pzed. No. 0.00036;			
Matches 31; Conservative	0; Mismatches 0; Indels 0; Gaps 0;			
Oy 259	GTGTCGAGTACAGCTGCCGAGAGTGACT 289			
Db 257	GTGTCGAGTACAGCTGCCGAGAGTGACT 287			
RESULT 9				
AZ876400	627 bp DNA linear GSS 21-FEB-2001			
LOCUS	2M0191D21F Mouse 10kb plasmid UBC1M library Mus musculus genomic			
DEFINITION	clone UBC2M0191D21 F, DNA sequence.			
ACCESSION	AZ876400			
VERSION	AZ876400.1 GI:13087273			
KEYWORDS	GSS.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Murinae; Mus. 1 (bases 1 to 627)			
AUTHORS	Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weis,R.			
TITLE	Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts			
JOURNAL	Unpublished (2000)			
COMMENT	Contact: Robert B. Weiss University of Utah Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA Tel: 801 585 5606 Fax: 801 585 7177 Email: ddunn@genetics.utah.edu Insert Length: 10000 Std Error: 0.00 Plate: 0191 row: D column: 21 Seq primer: CGTTGTAACGACGCGCAGT Class: plasmid ends High quality sequence stop: 627. Location/Qualifiers 1. .627 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /clone="UUC2M0191D21" /clone_1fb="Mouse 10kb plasmid UBC1M library" /sex="Male" /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-" /note="Vector: pMD24nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repeated with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The ligated DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD2 (gi 4732114 gb AF129072.1), a copy-number			

Seq primer: SP6
Class: BAC ends.

FEATURES

Location/Qualifiers

1..147
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_1ib="RPCI-24-79P16"
/sex="Male"
/cell_type="Spleen/Brain"
/note="Vector: pTRABACT; Site 1: BamHI, Site 2: BamHI, RPCI-24 Mouse BAC library produced by Pieter de Jong. The library was cloned in the pTRABACT cloning vector at the BamHI sites using MboI partially digested male C57BL/6J DNA."

BASE COUNT 30 a 44 c 45 g 28 t

ORIGIN

Query Match 3.4%; Score 26; DB 17; Length 147;

Best Local Similarity 100.0%; Pred. No. 0.077; Mismatches 0; Indels 0; Gaps 0;

Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 264 CGAGTACAGCTGCCGCGAGCTGCACT 289

DB 101 CGAGTACAGCTGCCGCGAGCTGCACT 126

RESULT 12
BG711211 368 bp mRNA linear EST 08-MAY-2001
p91n.pk006.114 Normalized Liver Library Gallus gallus cDNA clone
p91n.pk006.114 5' similar to g1485130 ref|NM_005195.1| Homo
sapiens CCAT/enhancer binding protein (C/EBP), delta (CEBPd),
mRNA, mRNA sequence.

ACCESSION BG711211 GI:14005161
VERSION BG711211.1
KEYWORDS EST.
SOURCE chicken.
ORGANISM Gallus gallus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianidae; Gallus.
1 (bases 1 to 368)
Burnside, J., Morgan, R.W. and Cogburn, L.A.
Chicken ESTs from a normalized liver library
Unpublished (2001)
Contact: Joan Burnside
Molecular Endocrinology
University of Delaware
40 Townsend Hall, Newark, DE 19717, USA
Tel: 302 831-1345
Fax: 302 831-3411
Email: joan@udel.edu, www.chickest.udel.edu.

FEATURES
source 1..368
Location/Qualifiers

1..368
/organism="Gallus gallus"
/db_xref="taxon:9031"
/clone="p91n.pk006.114"
/clone_1ib="Normalized Liver Library"
/sex="Male and Female"
/tissue_type="liver"
/lab_host="E.coli EMDH10B"
/note="Vector: PCMVSPORT 6"

BASE COUNT 74 a 127 c 102 g 65 t

ORIGIN

Query Match 3.3%; Score 25; DB 12; Length 368;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 618 GCCGCGGCCGCCGCCGCGAGCGCC 642

DB 233 GCCGCGGCCGCCGCCGCGAGCGCC 257

RESULT 13

LOCUS AW434705

DEFINITION UI-R-BJ0p-afz-c-02-0-UI.61 UI-R-BJ0p Rattus norvegicus cDNA clone

UI-R-BJ0p-afz-c-02-0-UI 3', mRNA sequence.

ACCESSION AW434705

VERSION AW434705.1 GI:6966012

KEYWORDS EST.

SOURCE Norway rat.

ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 385)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery

Genome Res. 6 (9), 791-806 (1996)

JOURNAL

MEDLINE

COMMENT

Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
Oligo-dt track not found. Not 1 site shown in beginning of sequence
is likely internal to the message. cDNA library preparation: M.B.
Soares Lab Clone distribution: clones will be available through
Research Genetics (www.resgen.com) The following repetitive
elements were found in this cDNA sequence: 21-41.
xgc_richlow_complexity
Seg primer: M13 forward
POLYA=NO.

FEATURES

source

1..385
Location/Qualifiers
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-BJ0p-afz-c-02-0-UI"
/clone_1ib="UI-R-BJ0p"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/note="Vector: p773D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-BJ0p
library is a subtracted library derived from the UI-R-A1,
UI-R-A1, UI-R-A1, UI-R-A1, UI-R-A1, UI-R-A1, and
UI-R-A1 libraries. These libraries represent tissues from
rat atrium at 16.5 dpc, ventricle at 16.5 dpc, AV canal
at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc, AV
canal at 15 dpc, and ventricle at 13 dpc. The tag is a
string of 5-6 nucleotides present between the Not I site
and the oligo-dt track. The library was constructed as
described by Bonaldo, Lennon and Soares, Genome Research
6: 791-806, 1996.

TAG_L1B=UI-R-BJ0p
TAG_TISSUE=AV canal at 16.5 dpc
TAG_SEQ=GAACCC

BASE COUNT 75 a 127 c 122 g 61 t

ORIGIN

Query Match 2.8%; Score 21; DB 10; Length 385;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 616 AAGCGCGGCCGCCGCCGCGCGG 636

DB 19 AAGCGCGGCCGCCGCCGCGG 39

[illegible]

```

SOURCE          human.
ORGANISM        Homo sapiens
REFERENCE       Bukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS         Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE           1 (bases 1 to 477)
JOURNAL         Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
MEDLINE         Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
COMMENT         Hood,L.
                Sequence-tagged connectors: A sequence approach to mapping and
                scanning the human genome
                Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
                99380589
                Contact: Mahairas GG, Wallace JC, Hood L
                High Throughput Sequencing Center
                University of Washington
                401 Queen Anne Avenue North, Seattle, WA 98109, USA
                Tel: (206) 616-3618
                Fax: (206) 616-3887
                Email: jwallace@u.washington.edu
                Clones may be purchased from Research Genetics (info@resgen.com).
                BAC end Web Server: http://www.htsc.washington.edu
                Plate: 3158 Row: N Column: 3
                Seq primer: T7
                Class: BAC ends
                High quality sequence stop: 477.
FEATURES
source          location/Qualifiers
                1..477
                /Organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="Plate=3158 Col=3 Row=N"
                /clone_1kb="CIT Approved Human Genomic Sperm Library D"
                /sex="male"
                /note="Organ: sperm; Vector: pBelobAC11; BAC clones in
                E-Coli DH10B"
BASE COUNT     76 a      171 c      143 g      82 t      5 others
ORIGIN
Query Match    2.8%; Score 21; DB 17; Length 477;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY             600 GCCGCGAAGGGCCGGAAGCC 620
                |||||
                |||||
Db              171 GCCGCAGAAGGGCCGGAAGCC 191

```


GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:06:06 ; Search time 53.2 Seconds
(without alignments) 475.895 Million cell updates/sec

Title: US-09-867-274-2
Perfect score: 1049

Sequence: 1 QGWAFFKDATEIIPELGEY.....KRPRAASAKAQAELENAY 190

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

Database :
A_GeneSeq_101002.*

1: /SIDS2/gcgdata/geneseq/genesseq-emb1/AAl980.DAT:*
2: /SIDS2/gcgdata/geneseq/genesseq-emb1/AAl981.DAT:*
3: /SIDS2/gcgdata/geneseq/genesseq-emb1/AAl982.DAT:*
4: /SIDS2/gcgdata/geneseq/genesseq-emb1/AAl983.DAT:*
5: /SIDS2/gcgdata/geneseq/genesseq-emb1/AAl984.DAT:*
6: /SIDS2/gcgdata/geneseq/genesseq-emb1/AAl985.DAT:*
7: /SIDS2/gcgdata/geneseq/genesseq-emb1/AAl986.DAT:*
8: /SIDS2/gcgdata/geneseq/genesseq-emb1/AAl987.DAT:*
9: /SIDS2/gcgdata/geneseq/genesseq-emb1/AAl988.DAT:*
10: /SIDS2/gcgdata/geneseq/genesseq-emb1/AAl989.DAT:*
11: /SIDS2/gcgdata/geneseq/genesseq-emb1/AAl990.DAT:*
12: /SIDS2/gcgdata/geneseq/genesseq-emb1/AAl991.DAT:*
13: /SIDS2/gcgdata/geneseq/genesseq-emb1/AAl992.DAT:*
14: /SIDS2/gcgdata/geneseq/genesseq-emb1/AAl993.DAT:*
15: /SIDS2/gcgdata/geneseq/genesseq-emb1/AAl994.DAT:*
16: /SIDS2/gcgdata/geneseq/genesseq-emb1/AAl995.DAT:*
17: /SIDS2/gcgdata/geneseq/genesseq-emb1/AAl996.DAT:*
18: /SIDS2/gcgdata/geneseq/genesseq-emb1/AAl997.DAT:*
19: /SIDS2/gcgdata/geneseq/genesseq-emb1/AAl998.DAT:*
20: /SIDS2/gcgdata/geneseq/genesseq-emb1/AAl999.DAT:*
21: /SIDS2/gcgdata/geneseq/genesseq-emb1/AAl2000.DAT:*
22: /SIDS2/gcgdata/geneseq/genesseq-emb1/AAl2001.DAT:*
23: /SIDS2/gcgdata/geneseq/genesseq-emb1/AAl2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1049	100.0	190	23	ABB07207	Human cloaked-2 po
2	1049	100.0	213	21	AAB26106	Human DAN/Cerberus
3	1049	100.0	213	21	AAV96429	Human TGF- β 2 bin
4	1049	100.0	213	22	AAV97589	Human TGF- β 2 bin
5	1049	100.0	213	23	ABG34061	Human Pro peptide
6	1049	100.0	213	23	ABB07209	Human cloaked-2 po
7	1049	100.0	213	23	AAE17089	Human osteolevin p
8	1040	99.1	213	21	AAV96430	Human TGF- β 2 bin
9	1040	99.1	213	21	AAV96436	Human TGF- β 2 bin
10	1033	98.5	213	21	AAV96431	Verret TGF- β 2 bin

11	974	32.9	213	21	AAV66103	Rat TGF- β binding
12	954	90.9	367	21	AAB36415	Human DNAN/Cerberus
13	951	90.7	188	23	AAB07208	Mouse cloaked-2 po
14	951	90.7	211	23	AAV96432	Murine TGF- β 2a b1
15	951	90.7	211	23	AAB07210	Mouse cloaked-2 po
16	901.5	85.9	176	21	AAV96434	Bovine TGF- β 2a b1
17	758	72.3	139	22	AAV73067	Human bone marrow
18	758	72.3	139	23	AABG42911	Human peptide encoc
19	344	32.8	206	21	AAV75981	Murine skin cell f
20	344	32.8	206	21	AAV76031	Murine skin cell f
21	344	32.8	206	22	AABS5920	Skin cell protein,
22	344	32.8	206	22	AABS5970	Skin cell protein,
23	344	32.8	206	23	AAB72120	Murine protein isc
24	344	32.8	206	23	AAB72120	Murine protein isc
25	341	32.5	206	18	AAV09408	Human small CCN-11
26	341	32.5	206	19	AAV08704	Human small CCN-11
27	341	32.5	206	20	AAV05711	Homo sapiens fetala
28	341	32.5	206	21	AAV10233	Human adult retina
29	341	32.5	206	23	AAB95457	Human angiogenesi
30	341	32.5	206	23	AAB09200	Human small CCN-11
31	341	32.5	206	23	AAB09200	SCGF CNM family p
32	341	32.5	183	23	AAB84851	Human PRO32 protei
33	337	32.1	183	23	AAB07213	Human cloaked-1 po
34	336.5	32.1	206	23	AAU82653	Human breast tumor
35	324.5	30.9	182	23	AAU82655	Human breast tumor
36	148	14.1	50	20	AAV12009	Human 5' EST secret
37	114	10.9	116	18	AAW47654	Secreted protein A
38	114	10.9	116	18	AAW47654	Human secreted prote
39	104.5	10.0	102	20	AAV03225	Amino acid sequenc
40	104.5	10.0	102	21	AAV10277	Human fetal kidney
41	95.5	9.1	418	22	AAAB3353	Human protein sequ
42	92	8.8	832	22	AABG3165	Novel human diagn
43	90.5	8.6	168	21	AAV04014	Amino acid sequenc
44	89.5	8.5	184	19	AAV09223	Human b57 protein
45	89.5	8.5	184	20	AAV42173	Human DMX protein

ALIGNMENTS

RESULT 1
ABB07207
ID ABB07207 standard; Protein; 190 AA

AC ABB07207

DT 26-MAR-2002 (first entry)
xx

Human cloaked-2 polypeptide mature protein sequence

KM cloaked-; cysteine knot motif; nephrotoxic; cardant; immunomodulator
KM hepatocellular; antinflammatory; anticholergic; cytosolic; neuroprotective
KM antiischemic; hypotensive; antiarrhythmic; antiatherosclerotic; muscular;
KM antidiabetic; anorectic; gene therapy; cell therapy; antisense therapy;
KM human.

OS Homo sapiens

FH	Key	Location/Qualifiers
FT	Miscellaneous	16

FT /note= "this region has been repeated twice in the
FT sequence provided in the sequence listing but has
FT been indicated correctly in the sequence in the
FT Figure"

PN WO200192308-A2.

PD 06-DEC-2001
xy

PF 29-MAY-2001; 2001WO-US17478.
XX

PR 01-JUN-2000; 2000US-208550P
PR 04-AUG-2000; 2000US-223542P

XX (AMGE-) AMGEN INC.
 PA Paszty CJ, Gao Y;
 PI WPI; 2002-114325/15.
 DR N-PSDB; ABA94293.
 XX
 PT New human and mouse cysteine-knot polypeptide designated as Cloaked-2,
 PT for treating or preventing kidney, heart (e.g. myocardial infarction)
 PT or liver (e.g. hepatitis) diseases
 PS
 PS Claim 13; Fig 1; 170pp; English.
 XX
 CC The invention relates to polypeptides comprising a cysteine knot motif
 CC and designated as Cloaked-2, derived from human and mouse. The cloaked-2
 CC polypeptides can be expressed by standard recombinant methodology. The
 CC cloaked-2 polynucleotides are useful in gene therapy and antisense
 CC therapy. The cloaked-2 polypeptides and polynucleotides are useful for
 CC treating, preventing, ameliorating or detecting diseases and disorders of
 CC the kidney (e.g. anemia, hypertension or low blood pressure), heart (e.g.
 CC cardiac hypertrophy, congestive heart failure, myocardial infarction,
 CC arrhythmias, atherosclerosis, hypertension or low blood pressure),
 CC skeletal muscle (e.g. muscular dystrophy or cachexia), placenta (e.g.
 CC congenital abnormalities or miscarriage), liver (e.g. hepatitis or
 CC cirrhosis), pancreas (e.g. diabetes or pancreatitis), thyroid (e.g.
 CC Grave's disease or myxedema) or adrenal cortex (e.g. Cushing's disease
 CC or Addison's disease), homeostasis or metabolic diseases (e.g. obesity,
 CC cancer or myopathies), infections, or autoimmune diseases. Selective
 CC binding agents may be used to modulate the biological activities of
 CC cloaked-2 polypeptides or to detect cloaked-2 polypeptide levels in a
 CC sample. Transgenic non-human animals are useful for drug candidate
 CC screening. The present sequence represents the human cloaked-2
 CC mature polypeptide sequence.
 CC
 SQ Sequence 190 AA;
 XX
 XX
 Query Match 100.0%; Score 1049; DB 23; Length 190;
 Best Local Similarity 100.0%; Pred. No. 1.5e-93;
 Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGMOAFKNDATETIPELGYPEPPELENKTNRAENGSRPHHPETKDVSEYSCREL 60
 DB 1 QGMOAFKNDATETIPELGYPEPPELENKTNRAENGSRPHHPETKDVSEYSCREL 60
 QY 61 HFTRYVTDPGCRSAKPVTELVCSGCGPARLLPNAIGRKWMPBSGDPFCIPDRYRAQR 120
 DB 61 HFTRYVTDPGCRSAKPVTELVCSGCGPARLLPNAIGRKWMPBSGDPFCIPDRYRAQR 120
 QY 121 VOLCPGGEAPRARKVRLVASCCKRKLTRFNQSELDKFGTEAARPKGRKPRPARSAK 180
 DB 121 VOLCPGGEAPRARKVRLVASCCKRKLTRFNQSELDKFGTEAARPKGRKPRPARSAK 180
 QY 181 ANQAELENAY 190
 DB 181 ANQAELENAY 190

RESULT 2
 ID AAB26106 standard; Protein; 213 AA.
 AC AAB26106;
 AC
 AC AAB26106;
 AC
 DT 15-JAN-2001 (first entry)
 XX
 XX Human DAN/Cerberus-related protein 6 (hDCR6) (exons 1 and 4).
 DE Human DAN/Cerberus-related protein 6 (hDCR6) (exons 1 and 4).
 XX
 XX Human; DNA/Cerberus-related protein 6; hDCR6; morphogenic protein;
 KM antagonist; BMP; cell growth; cell differentiation; bone formation;
 KW gene therapy.
 XX
 OS Homo sapiens.

XX W0200055193-A2.
 PN 21-SEP-2000.
 PD
 XX
 PF 02-MAR-2000; 2000WO-US05537.
 XX
 XX 12-MAR-1999; 99US-0124118.
 XX
 XX (REGGE-) REGENERON PHARM INC.
 PA
 XX
 PI Economides AN;
 PI WPI; 2000-638179/61.
 DR N-PSDB; AAA94051.
 XX
 XX Novel isolated, human DNA/Cerberus related protein 6 which include
 PT natural homologue, and polypeptides comprising DCR6 domain and nucleic
 PT acids encoding the proteins which are useful as probes and primers -
 XX
 PS Claim 8; Fig 3; 40pp; English.
 XX
 XX The present sequence comprises the amino acid sequence encoded by exons 1
 CC and 4 of the human DAN/Cerberus-related protein 6 (hDCR6) coding
 CC sequence. The coding sequence was isolated from a human kidney cDNA
 CC library containing exons 1 and 4 of the sequence. hDCR6 is closely
 CC related to the DAN and DCR5 proteins, both of which act as antagonists of
 CC morphogenic proteins such as BMP. It is possible that the hDCR6 gene and
 CC protein can be used as immunogens, modulators of cell function, growth
 CC and differentiation, to reduce undesirable bone formation, to identify
 CC DCR6 binding agents, in diagnosis, and in gene therapy.
 CC
 SQ Sequence 213 AA;
 XX
 XX
 Query Match 100.0%; Score 1049; DB 21; Length 213;
 Best Local Similarity 100.0%; Pred. No. 1.7e-93;
 Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGMOAFKNDATETIPELGYPEPPELENKTNRAENGSRPHHPETKDVSEYSCREL 60
 DB 24 QGMOAFKNDATETIPELGYPEPPELENKTNRAENGSRPHHPETKDVSEYSCREL 83
 QY 61 HFTRYVTDPGCRSAKPVTELVCSGCGPARLLPNAIGRKWMPBSGDPFCIPDRYRAQR 120
 DB 84 HFTRYVTDPGCRSAKPVTELVCSGCGPARLLPNAIGRKWMPBSGDPFCIPDRYRAQR 143
 QY 121 VOLCPGGEAPRARKVRLVASCCKRKLTRFNQSELDKFGTEAARPKGRKPRPARSAK 180
 DB 144 VOLCPGGEAPRARKVRLVASCCKRKLTRFNQSELDKFGTEAARPKGRKPRPARSAK 203
 QY 181 ANQAELENAY 190
 DB 204 ANQAELENAY 213

RESULT 3
 ID AAY96429 standard; Protein; 213 AA.
 AC AAY96429;
 AC
 AC AAY96429;
 AC
 DT 12-SEP-2000 (first entry)
 XX
 XX Human TGF-beta binding protein (BBER).
 DE Human TGF-beta binding protein (BBER).
 XX
 XX osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KW BBER; gene therapy; antisense therapy; fracture; bone mineralization.
 XX
 OS Homo sapiens.
 XX
 PN W0200032773-A1.
 XX
 PD 08-JUN-2000.

XX 24-NOV-1999; 99WO-US27990.
 XX 27-NOV-1998; 98US-0110283.
 XX (DARW-) DARWIN DISCOVERY LTD.
 PA
 XX Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepker BW;
 PI Van Ness J, Winkler DG;
 XX MPI; 2000-412321/35.
 DR N-PSDB; AAA29055.
 XX
 XX Nucleic acids (1) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures
 XX
 PS Claim 2; Page 116; 162pp; English.
 XX
 XX This shows the human transforming growth factor-beta (TGF-beta)
 CC binding protein designated hBEER. The cDNA and protein may be used for
 CC prevention, treatment and diagnosis of diseases associated with
 CC inappropriate BEER expression. For example, they may be used to treat
 CC disorders associated with decreased TGF-beta BP expression. The cDNA or
 CC vectors may be administered to treat diseases by rectifying mutations or
 CC deletions in a patient's genome that affect the activity of BEER by
 CC expressing inactive proteins or to supplement the patient's own production
 CC of BEER polypeptides. The nucleic acids may be used for recombinant
 CC production of BEER, gene therapy, antisense therapy, as probes for
 CC diagnostic assays and for functional studies. BEER may be used to raise
 CC antibodies and for identification of BEER modulators. BEER antagonists
 CC may be used to increase bone mineral content for the treatment of
 CC disorders such as osteopenia, osteoporosis, fractures and other
 CC disorders associated with low mineral content.
 XX
 SQ Sequence 213 AA;
 Query Match 100.0%; Score 1049; DB 21; Length 213;
 Best Local Similarity 100.0%; Pred. No. 1.7e-93;
 Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 OGMQAFKNDATETIIPELGEYEPPEPELENNTNRAENGRRPHHPFTKDVSEYSCREL 60
 DB 24 OGMQAFKNDATETIIPELGEYEPPEPELENNTNRAENGRRPHHPFTKDVSEYSCREL 83
 QY 61 HFTRYVTDGCRSAKPVTELVCSCGGCPARLLPNAIGRKWMPSPGDPFRCTPDYRAQR 120
 DB 84 HFTRYVTDGCRSAKPVTELVCSCGGCPARLLPNAIGRKWMPSPGDPFRCTPDYRAQR 143
 QY 121 VOLLCPGGEAPPAKRVRLVASCCKRLLTRFNQSELKDFTEARPPQKGRPRPARASAK 180
 DB 144 VOLLCPGGEAPPAKRVRLVASCCKRLLTRFNQSELKDFTEARPPQKGRPRPARASAK 203
 QY 181 ANQAELENAY 190
 DB 204 ANQAELENAY 213
 RESULT 4
 AA97589
 ID AA97589 standard; Protein; 213 AA.
 XX
 XX AA97589;
 DT 05-APR-2001 (first entry)
 XX
 XX Human secreted protein PRO7476.
 DE
 XX
 XX Secreted protein; human; PRO protein; neoplastic cell growth; tumour;
 KW proliferation; leukaemia; lymphoid malignancy; inflammatory disorder;
 KW angiogenic disorder; immunologic disorder; PRO7476.
 XX
 OS Homo sapiens.

XX WO200075317-A2.
 XX
 XX 14-DEC-2000.
 XX
 XX 15-MAY-2000; 2000WO-US13358.
 XX
 XX 09-JUN-1999; 99US-0138385.
 XX 20-JUL-1999; 99US-0144790.
 XX 03-AUG-1999; 99US-0146843.
 XX 10-AUG-1999; 99US-0148188.
 XX 17-AUG-1999; 99US-0149320.
 XX 17-AUG-1999; 99US-0149327.
 XX 17-AUG-1999; 99US-0149396.
 XX 20-AUG-1999; 99US-0150114.
 XX 31-AUG-1999; 99US-0151700.
 XX 31-AUG-1999; 99US-0151734.
 XX
 XX (GETH) GENENTECH INC.
 XX
 XX Botstein DA, Goddard A, Gurney AL, Smith V, Watanabe CK, Wood WI;
 PI MPI; 2001-071075/08.
 XX N-PSDB; AAA91023.
 XX
 XX Antibodies against PRO polypeptides, useful for diagnosing and treating
 PT tumours are associated with gene amplification, neoplastic cell growth
 PT and proliferation in mammals -
 XX
 PS Claim 61; Fig 20; 143pp; English.
 XX
 XX This sequence is a human PRO protein of the invention. The PRO
 CC proteins are secreted proteins. Antagonists or antibodies of PRO
 CC polypeptides are useful for diagnosing and treating tumours are
 CC associated with gene amplification, neoplastic cell growth and
 CC proliferation in mammals, and those conditions characterised by
 CC overexpression and/or activation of the amplified genes. Such conditions
 CC include benign or malignant tumours (e.g. renal, liver, kidney, bladder,
 CC breast, gastric, ovarian, colorectal, prostate, pancreatic, lung, vulval,
 CC thyroid, hepatic carcinomas, sarcomas, glioblastomas and various head and
 CC neck tumours), leukaemias and lymphoid malignancies, neuronal, glial,
 CC astrocytal, hypochalamic, and other glandular, macrophageal, epithelial,
 CC stromal and blastocoele disorders; and inflammatory, angiogenic and
 CC immunologic disorders. These may further be used to qualitatively or
 CC quantitatively detect the expression of proteins encoded by the
 CC amplified genes, and in tumour diagnostics or prognostics. The PRO
 CC polypeptide or its antagonist may be used for the preparation of a
 CC medicament in the treatment of a condition, which is responsive to the
 CC PRO polypeptide, its antagonist or anti-PRO antibody.
 XX
 SQ Sequence 213 AA;
 Query Match 100.0%; Score 1049; DB 22; Length 213;
 Best Local Similarity 100.0%; Pred. No. 1.7e-93;
 Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 OGMQAFKNDATETIIPELGEYEPPEPELENNTNRAENGRRPHHPFTKDVSEYSCREL 60
 DB 24 OGMQAFKNDATETIIPELGEYEPPEPELENNTNRAENGRRPHHPFTKDVSEYSCREL 83
 QY 61 HFTRYVTDGCRSAKPVTELVCSCGGCPARLLPNAIGRKWMPSPGDPFRCTPDYRAQR 120
 DB 84 HFTRYVTDGCRSAKPVTELVCSCGGCPARLLPNAIGRKWMPSPGDPFRCTPDYRAQR 143
 QY 121 VOLLCPGGEAPPAKRVRLVASCCKRLLTRFNQSELKDFTEARPPQKGRPRPARASAK 180
 DB 144 VOLLCPGGEAPPAKRVRLVASCCKRLLTRFNQSELKDFTEARPPQKGRPRPARASAK 203
 QY 181 ANQAELENAY 190
 DB 204 ANQAELENAY 213

RESULT 5
 ABG34061
 ID ABG34061 standard; Protein; 213 AA.
 XX
 AC ABG34061;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Human Pro peptide #32.
 XX
 KW Human; PRO; secreted protein; transmembrane protein;
 XX genetic disorder; tumour; cancer.
 XX
 OS Homo sapiens.
 XX
 WN WO200224888-A2.
 XX
 PD 28-MAR-2002.
 XX
 PF 29-AUG-2001; 2001WO-US27099.
 XX
 PR 01-SEP-2000; 2000US-229896P.
 PR 05-SEP-2000; 2000US-230621P.
 PR 22-SEP-2000; 2000US-235147P.
 PR 10-NOV-2000; 2000WO-US30873.
 PR 12-JAN-2001; 2001US-261878P.
 PR 16-JAN-2001; 2001US-261910P.
 PR 16-JAN-2001; 2001US-261939P.
 PR 25-JAN-2001; 2001US-264395P.
 PR 02-FEB-2001; 2001US-266421P.
 PR 09-FEB-2001; 2001US-267623P.
 PR 28-FEB-2001; 2001WO-US06520.
 PR 09-MAR-2001; 2001US-274399P.
 PR 03-APR-2001; 2001US-280982P.
 PR 04-APR-2001; 2001US-282129P.
 PR 04-APR-2001; 2001US-282199P.
 PR 09-MAY-2001; 2001US-290589P.
 PR 25-MAY-2001; 2001WO-US17092.
 PR 01-JUN-2001; 2001WO-US17800.
 PR 20-JUN-2001; 2001WO-US19692.
 PR 29-JUN-2001; 2001WO-US21066.
 PR 09-JUL-2001; 2001WO-US21735.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Baker KP, Eaton DL, Filvaroff E, Goddard A, Grimaldi JC;
 PI Guney AL, Smith V, Stephan J, Watanabe CK, Wood WI, Zhang Z;
 PI Fong S;
 XX
 DR WPI; 2002-362426/39.
 DR N-PSDB; ABR69992.
 XX
 PT New PRO polypeptides and polynucleotides encoding the polypeptides,
 PT useful in gene therapy, chromosome identification, tissue typing, or
 PT for genetic analysis of individuals with genetic disorders -
 XX
 PS Claim 11; Figure 64; 218pp; English.
 XX
 CC This invention relates to the cDNA and protein sequences of novel
 CC secreted and transmembrane polypeptides PRO polypeptides. The
 CC invention also comprises a method for producing the proteins of the
 CC invention by recombinant means and antibodies specific for the protein
 CC of the invention. The antibody may be used for detecting the PRO
 CC proteins of the invention and may be used to modify their activity.
 CC polynucleotides may be used as hybridisation probes for a cDNA library
 CC to isolate the full-length PRO cDNA or to isolate other cDNAs, to
 CC construct hybridisation probes for mapping the gene which encodes that
 CC PRO and for genetic analysis of individuals with genetic disorders, in
 CC assay to identify other proteins or molecules involved in binding
 CC reaction, to generate transgenic animals or knock-out animals which in
 CC turn are useful in the development and screening of therapeutically
 CC useful reagents, for chromosome identification, and tissue typing. The

Query Match	100.0%;	Score 1049;	DB 23;	Length 213;
Best Local Similarity	100.0%;	Pred. No. 1..7e-93;		
Matches 190;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
1	QGMWAFKNDTEILPELGEYRPPPELNNKTMRAENGGRPHHPETQDVSEYSCREL 60			
24	QGMWAFKNDTEILPELGEYRPPPELNNKTMRAENGGRPHHPETQDVSEYSCREL 83			
QY	61 HFTRYVTDGCRSAKVTETLVCSGCGPARLLPNAIGRKMRPSGDPFCIDRYRAQR 120			
DB	84 HFTRYVTDGCRSAKVTETLVCSGCGPARLLPNAIGRKMRPSGDPFCIDRYRAQR 143			
QY	121 VOLLCPGGEAPRARKVRLVASCCKRLTRFHNQSELKDFGTEARPDQGRKPPRRASAK 180			
DB	144 VOLLCPGGEAPRARKVRLVASCCKRLTRFHNQSELKDFGTEARPDQGRKPPRRASAK 203			
QY	181 ANQAELENAY 190			
DB	204 ANQAELENAY 213			
RESULT 6				
ABB07209				
1D	ABB07209 standard; Protein; 213 AA.			
XX	ABB07209;			
AC				
XX				
DT	26-MAR-2002 (first entry)			
XX				
DE	Human cloaked-2 polypeptide sequence.			
XX				
XX	Cloaked-2; cysteine knot motif; nephrotropic; cardiac; immunomodulator;			
XX	hepatotropic; antiinflammatory; antithyroid; cytostatic; neuroprotective;			
KW	antianemic; hypotensive; antiarrhythmic; antiarteriosclerotic; muscular;			
KW	antidiabetic; anorectic; gene therapy; cell therapy; antitense therapy;			
KM	human.			
XX				
OS	Homo sapiens.			
XX				
FH	Key	Location/Qualifiers		
FT	Peptide	1..23		
FT	/note= "signal peptide"			
FT	Protein	24..213		
FT	/note= "mature protein (ABB07207)"			
XX				
PN	MO200192308-A2.			
XX				
PD	06-DEC-2001.			
XX				
PF	29-MAY-2001; 2001WO-US17478.			
XX				
PR	01-JUN-2000; 2000US-208550P.			
PR	04-AUG-2000; 2000US-223542P.			
XX				
PA	(AMGE-) AMGEN INC.			
XX				
PI	Paszty CJ, Gao Y;			
XX				
DR	WPI; 2002-114325/15.			
XX				
DR	N-PSDB; ABA94293.			
XX				
PT	New human and mouse cystine-knot polypeptide designated as Cloaked-2,			
PT	for treating or preventing kidney, heart (e.g. myocardial infarction)			
PT	or liver (e.g. hepatitis) diseases -			

XX Example 1; Fig 1; 170pp; English.

CC The invention relates to polypeptides comprising a cysteine knot motif
 CC and designated as Cloaked-2, derived from human and mouse. The cloaked-2
 CC polypeptides can be expressed by standard recombinant methodology. The
 CC cloaked-2 polynucleotides are useful in gene therapy and antisense
 CC therapy. The cloaked-2 polypeptides and polynucleotides are useful for
 CC treating, preventing, ameliorating or detecting diseases and disorders of
 CC the kidney (e.g. anemia, hypertension or low blood pressure), heart (e.g.
 CC cardiac hypertrophy, congestive heart failure, myocardial infarction,
 CC arrhythmias, atherosclerosis, hypertension or low blood pressure),
 CC skeletal muscle (e.g. muscular dystrophy or cachexia), placenta (e.g.
 CC congenital abnormalities or miscarriage), liver (e.g. hepatitis or
 CC cirrhosis), pancreas (e.g. diabetes or pancreatitis), thyroid (e.g.
 CC Grave's disease or myxedema) or adrenal cortex (e.g. Cushing's disease
 CC or Addison's disease), osteostasis or metabolic diseases (e.g. obesity,
 CC cancer or myopathies), infections, or autoimmune diseases. Selective
 CC binding agents may be used to modulate the biological activities of
 CC cloaked-2 polypeptides or to detect cloaked-2 polypeptide levels in a
 CC sample. Transgenic non-human animals are useful for drug candidate
 CC screening. The present sequence represents the human cloaked-2
 CC polypeptide sequence.

CC Sequence 213 AA;

Query Match 100.0%; Score 1049; DB 23; Length 213;
 Best Local Similarity 100.0%; Pred. No. 1.7e-93;
 Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGWAQFNDAATEIIPELGEYEPPEPELENNKTNRRAENGSRPHHPFETKDVSEYSCREL 60
 DB 24 QGWAQFNDAATEIIPELGEYEPPEPELENNKTNRRAENGSRPHHPFETKDVSEYSCREL 83
 QY 61 HFTRYVTDGPCRSAPVTELVCSGGCGPARLLPNAIGRKMMWRPSGDPFRICIPRYRAQR 120
 DB 84 HFTRYVTDGPCRSAPVTELVCSGGCGPARLLPNAIGRKMMWRPSGDPFRICIPRYRAQR 143
 QY 121 VOLLCRGSEAPRARKVRLVASCCKRLTRFNQSELDKGTGEARPOKGRPRPARSAK 180
 DB 144 VOLLCRGSEAPRARKVRLVASCCKRLTRFNQSELDKGTGEARPOKGRPRPARSAK 203

QY 181 ANQAELENAY 190
 DB 204 ANQAELENAY 213

RESULT 7
 AAE17089
 ID AAE17089 standard; Protein; 213 AA.

XX AAE17089;

XX 18-APR-2002 (first entry)

DE Human osteolevin protein.

XX Human; osteolevin; osteopathic; cytoskeletal; bone formation; osteoporosis;
 KM Van Buchem-sclerosteosis disease; sclerosteosis; transgenic animal;
 KW Paget's disease; chromosome 17.

XX Homo sapiens.

OS Key Location/Qualifiers

PH Peptide 1..19 /label= Signal_peptide

FT Protein 20..213 /label= Mature_osteolevin_protein

FT Misc-difference 10 /note= "During polymorphism wild type Val is
 FT substituted with Ile"
 XX
 PN

WO200198491-A2.

XX 27-DEC-2001.

XX 15-JUN-2001; 2001WO-EP06795.

XX 19-JUN-2000; 2000EP-0112867.

PA (HOFF) HOFFMANN LA ROCHE & CO AG F.
 PA (UWIN-) UNIV INGELING AMTWERPEN UTA.

PI Ballemans W, Ebeling M, Foerzler D, Patel N, Van Hul W,
 PI Vickers BH;

DR WPI: 2002-139789/18.
 DR N-PSDB: AAD27576, AAD27577.

PT Novel genetic polymorphisms in the Van Buchem-sclerosteosis disease
 PT region that are associated with abnormal bone formation useful for
 PT diagnosis and assessment of osteoporosis or sclerosteosis in humans
 XX
 PS Claim 11; Fig 4; 70pp; English.

CC The invention relates to a nucleic acid encoding osteolevin region
 CC polymorphisms. The invention also relates to genetic polymorphisms in
 CC the Van Buchem-sclerosteosis disease region that are associated with
 CC disorders resulting in either net excess bone formation or insufficient
 CC bone formation in humans. Osteolevin DNA is useful for screening for
 CC osteolevin polymorphisms associated with abnormal bone formation in a
 CC subject and for the presence of a heritably linked form of abnormal bone
 CC formation in a subject, by determining the presence of a polymorphism in
 CC the osteolevin nucleic acid sequence obtained from the subject.
 CC Osteolevin protein is useful for treating diseases associated with
 CC abnormal bone formation, such as sclerosteosis, Van Buchem's disease and
 CC Paget's disease. Nucleic acids which encode genes in the osteolevin
 CC region or their modified forms can also be used to generate either
 CC transgenic animals or knockout animals which are useful in the screening
 CC and development of therapeutically useful reagents. Osteolevin proteins
 CC are useful in pharmacological characterisation of novel modulators of the
 CC activity of protein and protein complexes. Human osteolevin gene is
 CC located on chromosome 17. The present sequence is human osteolevin
 CC protein.

CC Sequence 213 AA;

Query Match 100.0%; Score 1049; DB 23; Length 213;
 Best Local Similarity 100.0%; Pred. No. 1.7e-93;
 Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QGWAQFNDAATEIIPELGEYEPPEPELENNKTNRRAENGSRPHHPFETKDVSEYSCREL 60
 DB 24 QGWAQFNDAATEIIPELGEYEPPEPELENNKTNRRAENGSRPHHPFETKDVSEYSCREL 83

QY 61 HFTRYVTDGPCRSAPVTELVCSGGCGPARLLPNAIGRKMMWRPSGDPFRICIPRYRAQR 120

DB 84 HFTRYVTDGPCRSAPVTELVCSGGCGPARLLPNAIGRKMMWRPSGDPFRICIPRYRAQR 143

QY 121 VOLLCRGSEAPRARKVRLVASCCKRLTRFNQSELDKGTGEARPOKGRPRPARSAK 180

DB 144 VOLLCRGSEAPRARKVRLVASCCKRLTRFNQSELDKGTGEARPOKGRPRPARSAK 203

QY 181 ANQAELENAY 190

DB 204 ANQAELENAY 213

RESULT 8

AAV96430
 ID AAV96430 standard; Protein; 213 AA.

XX AAV96430;

XX 12-SEP-2000 (first entry)

DE Human TGF-beta binding protein (BEER) variant V101.
 XX osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KM BEER; variant; V101; gene therapy; antisense therapy; fracture;
 KM chromosome 17q12-21; bone mineralization.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 10 /label= V101
 FT /note= "wild type valine has been substituted with
 FT isoleucine"
 XX
 XX MO200032773-A1.
 XX
 PD 08-JUN-2000.
 XX
 XX 24-NOV-1999; 99WO-US27990.
 PF
 XX 27-NOV-1998; 98US-0110283.
 PR
 XX (DARW-) DARWIN DISCOVERY LTD.
 PA
 XX Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;
 PI Van Ness J, Winkler DG;
 XX
 DR N-PSDB; AAA29056.
 XX
 XX Nucleic acids (1) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures
 XX
 PS Claim 3; Page 119-120; 162pp; English.
 XX
 XX This shows a variant human transforming growth factor-beta (TGF-beta)
 CC binding protein designated BEER V101, which comprises a substitution of
 CC isoleucine for the wild-type valine at residue 10. The CDNA and protein
 CC may be used for prevention, treatment and diagnosis of diseases
 CC associated with inappropriate BEER expression. For example, they may be
 CC used to treat disorders associated with decreased TGF-beta BP expression.
 CC The CDNA or vectors may be administered to treat diseases by rectifying
 CC mutations or deletions in a patient's genome that affect the activity of
 CC BEER by expressing inactive proteins or to supplement the patients own
 CC production of BEER polypeptides. The nucleic acids may be used for
 CC recombinant production of BEER, gene therapy, antisense therapy, as
 CC probes for diagnostic assays and for functional studies. BEER may be used
 CC to raise antibodies and for identification of BEER modulators. BEER
 CC antagonists may be used to increase bone mineral content for the
 CC treatment of disorders such as osteopenia, osteoporosis, fractures and
 CC other disorders associated with low mineral content.
 XX
 SQ Sequence 213 AA;
 XX
 XX Query Match 99.1%; Score 1040; DB 21; Length 213;
 XX Best Local Similarity 99.5%; Pred. No. 1.3e-92;
 XX Matches 189; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 XX
 QY 1 QGWAQKNDATETIPELGEYEPPEPELENNKTNRAENGCRPPHPETDVSYSREL 60
 DB 24 QGWAQKNDATETIPELGEYEPPEPELENNKTNRAENGCRPPHPETDVSYSREL 83
 QY 61 HFTTYVVDGPGCRSAKPTTELVCSGCCGPARLLPVAIGKMMWRPSGDPFCIDRYAQR 120
 DB 84 HFTTYVVDGPGCRSAKPTTELVCSGCCGPARLLPVAIGKMMWRPSGDPFCIDRYAQR 143
 QY 121 VOLLCPGGEAPRARKVRLVASCCKRLTRPHNOSKLDFTGEARPOKGRKPPRARSASAK 180
 DB 144 VOLLCPGGEAPRARKVRLVASCCKRLTRPHNOSKLDFTGEARPOKGRKPPRARSASAK 203
 QY 181 ANQAELENAY 190
 |||||||||

DB 204 ANQAELENAY 213
 XX
 XX RESULT 9
 XX ID AAY96436
 XX AAY96436 standard; Protein; 213 AA.
 XX
 XX AAY96436;
 XX
 XX 12-SEP-2000 (first entry)
 XX
 XX Human TGF-beta binding protein (BEER) variant P38R.
 DE
 XX osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KM BEER; variant; P38R; gene therapy; antisense therapy; fracture;
 KM chromosome 17q12-21; bone mineralization.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 38 /label= P38R
 FT /note= "wild type proline has been substituted with
 FT arginine"
 XX
 XX MO200032773-A1.
 XX
 PD 08-JUN-2000.
 XX
 XX 24-NOV-1999; 99WO-US27990.
 PF
 XX 27-NOV-1998; 98US-0110283.
 PR
 XX (DARW-) DARWIN DISCOVERY LTD.
 PA
 XX Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;
 PI Van Ness J, Winkler DG;
 XX
 DR N-PSDB; AAA29062.
 XX
 XX Nucleic acids (1) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures
 XX
 PS Disclosure; Page 121; 162pp; English.
 XX
 XX This shows a variant human transforming growth factor-beta
 CC (TGF-beta) binding protein designated BEER P38R. The encoded protein
 CC comprises a substitution of arginine for the wild-type proline at
 CC residue 38. The CDNA and protein may be used for prevention, treatment
 CC and diagnosis of diseases associated with inappropriate BEER expression.
 CC For example, they may be used to treat disorders associated with
 CC decreased TGF-beta BP expression. The CDNA or vectors may be administered
 CC to treat diseases by rectifying mutations or deletions in a patient's
 CC genome that affect the activity of BEER by expressing inactive proteins
 CC or to supplement the patients own production of BEER polypeptides. The
 CC nucleic acids may be used for recombinant production of BEER, gene
 CC therapy, antisense therapy, as probes for diagnostic assays and for
 CC functional studies. BEER may be used to raise antibodies and for
 CC identification of BEER modulators. BEER antagonists may be used to
 CC increase bone mineral content for the treatment of disorders such as
 CC osteopenia, osteoporosis, fractures and other disorders associated with
 CC low mineral content.
 XX
 SQ Sequence 213 AA;
 XX
 XX Query Match 99.1%; Score 1040; DB 21; Length 213;
 XX Best Local Similarity 99.5%; Pred. No. 1.3e-92;
 XX Matches 189; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 XX
 QY 1 QGWAQKNDATETIPELGEYEPPEPELENNKTNRAENGCRPPHPETDVSYSREL 60
 |||||||||

DB 24 QGWAQKNDATETIIRLEGEYPEPPPELENNKTMRAENGRRPHHPFETKDVSEYSCEEL 83
 QY 61 HFTRYVTDGPCRSAPVTELVCSGGCGPARLLPNAIGRGKMWRRPSGDPFCIPDRYRAOR 120
 DB 84 HFTRYVTDGPCRSAPVTELVCSGGCGPARLLPNAIGRGKMWRRPSGDPFCIPDRYRAOR 143
 QY 121 VOLLCPGGAAPPAARVRLVASCCKRLTRFNQSELDKDFTEAARPOKGRPRPARSAK 180
 DB 144 VOLLCPGGAAPPAARVRLVASCCKRLTRFNQSELDKDFTEAARPOKGRPRPARSAK 203
 QY 181 ANQAELENNAY 190
 DB 204 ANQAELENNAY 213

RESULT 10
 ID AAY96431 standard; Protein, 213 AA.
 AC AAY96431;

DT 12-SEP-2000 (first entry)

DE Vervet TGF-beta binding protein (BEER).

KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 BEER; gene therapy; antisense therapy; fracture; bone mineralization.

OS Cercopithecus pygerythrus.

XX WO200032773-A1.

XX 08-JUN-2000.

XX 24-NOV-1999; 99WO-US27990.

XX 27-NOV-1998; 98US-0110283.

PA (DARW-) DARWIN DISCOVERY LTD.

PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;
 PI Van Ness J, Winkler DG;

DR WPI; 2000-412321/35.
 DR N-PSDB; AAA29057.

PT Nucleic acids (I) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures

PS Claim 4; Page 122-123; 162pp; English.

XX This shows a vervet transforming growth factor-beta (TGF-beta)
 CC binding protein designated VBER. The cDNA and protein may be used for
 CC prevention, treatment and diagnosis of diseases associated with
 CC inappropriate BEER expression. For example, they may be used to treat
 CC disorders associated with decreased TGF-beta BP expression. The cDNA or
 CC vectors may be administered to treat diseases by rectifying mutations or
 CC deletions in a patient's genome that affect the activity of BEER by
 CC expressing inactive proteins or to supplement the patients own production
 CC of BEER polypeptides. The nucleic acids may be used for recombinant
 CC production of BEER, gene therapy, antisense therapy, as probes for
 CC diagnostic assays and for functional studies. BEER may be used to raise
 CC antibodies and for identification of BEER modulators. BEER antagonists
 CC may be used to increase bone mineral content for the treatment of
 CC disorders such as osteopenia, osteoporosis, fractures and other disorders
 CC associated with low mineral content.

XX Sequence 213 AA;

Query Match 98.5%; Score 1033; DB 21; Length 213;
 Best Local Similarity 98.4%; Pred. No. 6.2e-92;
 Matches 187; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 QGWAQKNDATETIIRLEGEYPEPPPELENNKTMRAENGRRPHHPFETKDVSEYSCEEL 60
 DB 24 QGWAQKNDATETIIRLEGEYPEPPPELENNKTMRAENGRRPHHPFETKDVSEYSCEEL 83
 QY 61 HFTRYVTDGPCRSAPVTELVCSGGCGPARLLPNAIGRGKMWRRPSGDPFCIPDRYRAOR 120
 DB 84 HFTRYVTDGPCRSAPVTELVCSGGCGPARLLPNAIGRGKMWRRPSGDPFCIPDRYRAOR 143
 QY 121 VOLLCPGGAAPPAARVRLVASCCKRLTRFNQSELDKDFTEAARPOKGRPRPARSAK 180
 DB 144 VOLLCPGGAAPPAARVRLVASCCKRLTRFNQSELDKDFTEAARPOKGRPRPARSAK 203
 QY 181 ANQAELENNAY 190
 DB 204 ANQAELENNAY 213

RESULT 11
 ID AAY96433 standard; Protein, 213 AA.
 AC AAY96433;

DT 12-SEP-2000 (first entry)

DE Rat TGF-beta binding protein (BEER).

KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 BEER; gene therapy; antisense therapy; fracture; bone mineralization.

OS Rattus norvegicus.

XX WO200032773-A1.

XX 08-JUN-2000.

XX 24-NOV-1999; 99WO-US27990.

XX 27-NOV-1998; 98US-0110283.

PA (DARW-) DARWIN DISCOVERY LTD.

PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;
 PI Van Ness J, Winkler DG;

DR WPI; 2000-412321/35.
 DR N-PSDB; AAA29059.

PT Nucleic acids (I) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures

PS Claim 6; Page 125-126; 162pp; English.

XX This shows a rat transforming growth factor-beta (TGF-beta) binding
 CC protein designated RBER. The cDNA and protein may be used for
 CC prevention, treatment and diagnosis of diseases associated with
 CC inappropriate BEER expression. For example, they may be used to treat
 CC disorders associated with decreased TGF-beta BP expression. The cDNA or
 CC vectors may be administered to treat diseases by rectifying mutations or
 CC deletions in a patient's genome that affect the activity of BEER by
 CC expressing inactive proteins or to supplement the patients own production
 CC of BEER polypeptides. The nucleic acids may be used for recombinant
 CC production of BEER, gene therapy, antisense therapy, as probes for
 CC diagnostic assays and for functional studies. BEER may be used to raise
 CC antibodies and for identification of BEER modulators. BEER antagonists
 CC may be used to increase bone mineral content for the treatment of
 CC disorders such as osteopenia, osteoporosis, fractures and other disorders
 CC associated with low mineral content.

XX Sequence 213 AA;

Query Match 92.9%; Score 974; DB 21; Length 213;
 Best Local Similarity 92.1%; Pred. No. 3.2e-86;
 Matches 175; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 QGWAFFKNDATETIIPELGEYEPPEPELENNKTNNRAENGRRPHHPETKDVSEYSCREL 60
 DB 24 QGWAFFKNDATETIIPELGEYEPPEPELENNKTNNRAENGRRPHHPETKDVSEYSCREL 83
 QY 61 HPTTYVTGDCRSKAPVTELVCSGCCGPARLLPNAIGRGKWRPSGDPFCIDRRAQR 120
 DB 84 HYTFVTDGFCRSKAPVTELVCSGCCGPARLLPNAIGRGKWRPSGDPFCIDRRAQR 143
 QY 121 VOLLCPGGAAPRAKRVLVASCKCKRLTRFNQSELKDPGTEARPOGKRKPRPARSAK 180
 DB 144 VOLLCPGGAAPRAKRVLVASCKCKRLTRFNQSELKDPGTEARPOGKRKPRPARSAK 203
 QY 181 ANQAELENAY 190
 DB 204 ANQAELENAY 213

RESULT 12
 AAB26105
 ID AAB26105 standard; Protein; 367 AA.
 AC AAB26105;
 XX
 DT 15-JAN-2001 (first entry)
 XX
 DE Human DAN/Cerberus-related protein 6 (hDCR6) #1.
 XX
 KW Human; DNA/Cerberus-related protein 6; hDCR6; morphogenic protein;
 KM antagonist; BMP; cell growth; cell differentiation; bone formation;
 KW gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN MO200055193-A2.
 XX
 PD 21-SEP-2000.
 XX
 PF 02-MAR-2000; 2000WO-US05537.
 XX
 PR 12-MAR-1999; 99US-0124118.
 XX
 PA (REGG-) REGENERON PHARM INC.
 XX
 PI Economides AM;
 XX
 DR WPI; 2000-638179/61.
 DR N-PSDB; AAA94049, AAA94050.
 XX
 PT Novel isolated, human DNA/Cerberus related protein 6 which include
 PT natural homologue, and polypeptides comprising DCR6 domain and nucleic
 PT acids encoding the proteins which are useful as probes and primers -
 XX
 PS Claim 7; Fig 2; 40pp; English.
 XX
 CC The present sequence comprises the human DAN/Cerberus-related protein 6
 CC (hDCR6). Its coding sequence was isolated from a genomic DNA clone
 CC following identification using computer-based "virtual cloning". hDCR6
 CC is closely related to the DAN and DCR5 proteins, both of which act as
 CC antagonists of morphogenic proteins such as BMP. It is possible that
 CC the hDCR6 gene and protein can be used as immunogens, modulators of cell
 CC function, growth and differentiation, to reduce undesirable bone
 CC formation, to identify DCR6 binding agents, in diagnosis, and in gene
 CC therapy.
 CC
 SQ Sequence 367 AA;

Query Match 90.9%; Score 954; DB 21; Length 367;
 Best Local Similarity 54.9%; Pred. No. 5.2e-84;
 Matches 189; Conservative 0; Mismatches 1; Indels 154; Gaps 1;

QY 1 QGWAFFKNDATETIIPELGEYEPPEPELENNKTNNRAENGRRPHHPETKDVSEYSCREL 51
 DB 24 QGWAFFKNDATETIIPELGEYEPPEPELENNKTNNRAENGRRPHHPETKDVSEYSCREL 83
 QY 52 ----- 51
 DB 84 VGEWLAGAAAFRCRCGRMOOTLVRAQREDAGVVRVWHQGITRTSGAQKRGKFPSPG 143
 QY 52 ----- 51
 DB 144 NIGATSSCWTHWEGNKVAEKSTAQAPQPPPHNLGLWAKKGCNMGOGPPSVKME 203
 QY 52 -----VSEYSCRELHPTTYVTGDCRSKAPVTELVCSGCC 86
 DB 204 DKGASPHSPSRNGQLGSSDCLKGVSYSCRELHPTTYVTGDCRSKAPVTELVCSGCC 263
 QY 87 GPARLLPNAIGRGKWRPSGDPFCIDRRAQRVOLLCPGGAAPRAKRVLVASCKCKR 146
 DB 264 GPARLLPNAIGRGKWRPSGDPFCIDRRAQRVOLLCPGGAAPRAKRVLVASCKCKR 323
 QY 147 LTRFNQSELKDPGTEARPOGKRKPRPARSAKANOAELENAY 190
 DB 324 LTRFNQSELKDPGTEARPOGKRKPRPARSAKANOAELENAY 367

RESULT 13
 ABB07208
 ID ABB07208 standard; Protein; 188 AA.
 AC ABB07208;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Mouse cloaked-2 polypeptide mature protein sequence.
 XX
 KW Cloaked-2; cysteine knot motif; nephrotropic; cardiant; immunomodulator;
 KM hepatocytic; antiinflammatory; antithyroid; cytostatic; neuroprotective;
 KM antianemic; hypotensive; antiarrhythmic; antiarteriosclerotic; muscular;
 KM antidiabetic; anorectic; gene therapy; cell therapy; antisense therapy;
 KM mouse.
 XX
 OS Mus musculus.
 XX
 PN MO200192308-A2.
 XX
 PD 06-DEC-2001.
 XX
 PR 29-MAY-2001; 2001WO-US17478.
 XX
 PF 01-JUN-2000; 2000US-208550P.
 PR 04-AUG-2000; 2000US-223542P.
 XX
 PA (AMGE-) AMGEN INC.
 XX
 PI Paszty CJ, Gao Y;
 XX
 DR WPI; 2002-114325/15.
 DR N-PSDB; ABA94294.
 XX
 PT New human and mouse cysteine-knot polypeptide designated as Cloaked-2,
 PT for treating or preventing kidney, heart (e.g. myocardial infarction)
 PT or liver (e.g. hepatitis) diseases -
 XX
 PS Claim 13; Fig 2; 170pp; English.
 XX
 CC The invention relates to polypeptides comprising a cysteine knot motif

Db 432 GGGCCGACTTCGCTGCATCCCGACCGCTACCGCGCGAGCGCTGCTGTGT 491
 Qy 481 CCGGTGTGTAGGCGCGCGCGCGCGAAGTGTGCTGTGTGTGTGTGTGTGTGTGT 540
 Db 492 CCGGTGTGTAGGCGCGCGCGCGCGAAGTGTGCTGTGTGTGTGTGTGTGTGTGT 551
 Qy 541 CCGCTCACCGCTTCCACCAACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
 Db 552 CCGCTCACCGCTTCCACCAACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 611
 Qy 601 CCGCAAGAGGCGCGAAGCGCGCGCGCGCGAAGCGCGCGCGCGCGCGCGCGCG 660
 Db 612 CCGCAAGAGGCGCGAAGCGCGCGCGCGCGAAGCGCGCGCGCGCGCGCGCGCG 671
 Qy 661 CCGCAAGAGGCGCGAAGCGCGCGCGCGCGAAGCGCGCGCGCGCGCGCGCGCG 720
 Db 672 CCGCAAGAGGCGCGAAGCGCGCGCGCGCGAAGCGCGCGCGCGCGCGCGCGCG 731
 Qy 721 GAAACCG 759
 Db 732 GAAACCG 770

RESULT 5
 AAA91023
 ID AAA91023 standard; DNA; 2329 BP.

AC AAA91023;

DT 05-APR-2001 (first entry)

DE Human secreted protein PRO7476 coding sequence.

XX Secreted protein: human; PRO protein; neoplastic cell growth; tumour;
 KM proliferation; leukaemia; lymphoid malignancy; inflammatory disorder;
 KM angiogenic disorder; immunologic disorder; PRO7476; ds.

OS Homo sapiens.

XX Key Location/Qualifiers
 FT CDS 62..703
 FT /tag= a
 FT /product= PRO7476

PN WO200075317-A2.

PD 14-DEC-2000.

PF 15-MAY-2000; 2000WO-US13358.

PR 09-JUN-1999; 99US-0138385.

PR 20-JUL-1999; 99US-0144790.

PR 03-AUG-1999; 99US-0146843.

PR 10-AUG-1999; 99US-0148188.

PR 17-AUG-1999; 99US-0149320.

PR 17-AUG-1999; 99US-0149327.

PR 20-AUG-1999; 99US-0150114.

PR 31-AUG-1999; 99US-0151700.

PR 31-AUG-1999; 99US-0151734.

PA (GETH) GENENTECH INC.

PI Botstein DA, Goddard A, Gurney AL, Smith V, Watanabe CK, Wood WI;
 DR WPI, 2001-071075/08.
 DR P-PSDB; AAY97589.

XX Antibodies against PRO polypeptides, useful for diagnosing and treating
 PT tumors are associated with gene amplification, neoplastic cell growth
 PT and proliferation in mammals -
 PS Claim 50; Fig 19, 143p; English.

XX This sequence encodes a human PRO protein of the invention. The PRO
 CC proteins are secreted proteins. Antagonists or antibodies of PRO
 CC polypeptides are useful for diagnosing and treating tumors are
 CC associated with gene amplification, neoplastic cell growth and
 CC proliferation in mammals, and those conditions characterized by
 CC overexpression and/or activation of the amplified genes. Such conditions
 CC include benign or malignant tumors (e.g. renal, liver, kidney, bladder,
 CC breast, gastric, ovarian, colorectal, prostate, pancreatic, lung, uterine,
 CC thyroid, hepatic carcinomas, sarcomas, glioblastomas and various head and
 CC neck tumors); leukaemias and lymphoid malignancies; neuronal, glial,
 CC astrocytic, hypothalamic, and other glandular, macrophageal, epithelial,
 CC stromal and blastocoele disorders; and inflammatory, angiogenic and
 CC immunologic disorders. These may further be used to qualitatively or
 CC quantitatively detect the expression of proteins encoded by the
 CC amplified genes, and in tumour diagnostics or prognostics. The PRO
 CC polypeptide or its antagonist may be used for the preparation of a
 CC medicament in the treatment of a condition, which is responsive to the
 CC PRO polypeptide, its antagonist or anti-PRO antibody.

Sequence 2329 BP; 587 A; 645 C; 612 G; 485 T; 0 other;

Query Match 99.7%; Score 757; DB 22; Length 2329;
 Best Local Similarity 100.0%; Pred. No. 1.e-137;
 Matches 757; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 CTGGAAGGTGGGCGTCCCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 62
 Db 28 CTGGAAGGTGGGCGTCCCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 87
 Qy 63 CGTCTGCTGT 122
 Db 88 CGTGTGCTGT 147
 Qy 123 CAAGATATATGTCACAGAAATATATATATATATATATATATATATATATATAT 182
 Db 148 CAAGATATATGTCACAGAAATATATATATATATATATATATATATATATATAT 207
 Qy 183 GCTGTGGAACAACAAGACATATGAAACCGGCGGAGAGAGAGAGAGAGAGAGAGAG 242
 Db 208 GCTGTGGAACAACAAGACATATGAAACCGGCGGAGAGAGAGAGAGAGAGAGAGAG 267
 Qy 243 CTTTGAGACCAAGAAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 302
 Db 268 CTTTGAGACCAAGAAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 327
 Qy 303 GACCGATGAGCGCGT 362
 Db 328 GACCGATGAGCGCGT 387
 Qy 363 CGGCGCGGCGCGCTGT 422
 Db 388 CGGCGCGGCGCGCTGT 447
 Qy 423 GCCCGACTTTCGCTGT 482
 Db 448 GCCCGACTTTCGCTGT 507
 Qy 483 CGGT 542
 Db 508 CGGT 567
 Qy 543 CCTCACCGGCTTCCCAACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 602
 Db 568 CCTCACCGGCTTCCCAACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 627
 Qy 603 GCAGAGAGGCGGAAAGCGCGCGCGCGCGCGAGAGCGCGAGAGCGAGAGAGAGAG 662
 Db 628 GCAGAGAGGCGGAAAGCGCGCGCGCGCGAGAGCGCGAGAGCGAGAGAGAGAGAG 687
 Qy 663 GGAAGAGCGCTTCTAGAGAGCGCGCGCGCGCGCGAGAGCGAGAGAGAGAGAGAG 722
 Db 688 GGAAGAGCGCTTCTAGAGAGCGCGCGCGCGCGCGAGAGCGAGAGAGAGAGAGAG 747

ID	AA029056	standard; cDNA; 2301 BP.
XX	AAA29056	
AC	AAA29056;	
DT	12-SEP-2000	(first entry)
XX		
DE	Human TGF-beta binding protein (BEER) variant V101 cDNA.	
KW	osteopathic; transforming growth factor-beta; TGF-beta; binding protein;	
KM	BEER; variant; V101; gene therapy; antisense therapy; fracture;	
XX	chromosome 1/q12-21; bone mineralization; ss.	
OS	Homo sapiens.	
XX		
PH	Key	Location/Qualifiers
FT	CDS	48..689
FT		/*tag= a
FT		/label= BEER variant_V101
XX		/product= TGF-beta_binding_protein
PN	MO200032773-A1.	
XX		
PD	08-JUN-2000.	
PF	24-NOV-1999;	99WO-US27990.
XX		
PR	27-NOV-1998;	98US-0110283.
XX		
PA	(DARW-) DARWIN DISCOVERY LTD.	
XX		
PI	Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;	
PI	Van Ness J, Winkler DG;	
XX		
XX	WPI; 2000-412321/35.	
DR	P-PSDB; AAY96430.	
XX		
PT	Nucleic acid (1) encoding a transforming growth factor beta binding	
PT	protein, useful for identifying agents for treating osteopenia,	
PT	osteoporosis and fractures	
XX		
PS	Claim 1; Page 118-119; 162pp; English.	
XX		
CC	This cDNA encodes a variant human transforming growth factor-beta	
CC	(TGF-beta) binding protein designated BEER V101. The encoded protein	
CC	comprises a substitution of isoleucine for the wild-type valine at	
CC	residue 10. The cDNA and protein may be used for prevention, treatment	
CC	and diagnosis of diseases associated with inappropriate BEER expression.	
CC	For example, they may be used to treat disorders associated with	
CC	decreased TGF-beta BP expression. The cDNA or vectors may be administered	
CC	to treat diseases by rectifying mutations or deletions in a patient's	
CC	genome that affect the activity of BEER by expressing inactive proteins	
CC	or to supplement the patient's own production of BEER polypeptides. The	
CC	nucleic acids may be used for recombinant production of BEER, gene	
CC	therapy, antisense therapy, as probes for diagnostic assays and for	
CC	functional studies. BEER may be used to raise antibodies and for	
CC	identification of BEER modulators. BEER antagonists may be used to	
CC	increase bone mineral content for the treatment of disorders such as	
CC	osteopenia, osteoporosis, fractures and other disorders associated with	
CC	low mineral content.	
XX		
XX		
SO	Sequence 2301 BP; 569 A; 634 C; 614 G; 484 T; 0 other;	
XX		
Query Match	99.6%;	Score 755.8; DB 21; Length 2301;
Best Local Similarity	99.7%;	Prod. No. 1.8e-13;
Matches 757; Conservative	0; Mismatches 2;	Indels 0; Gaps 0;
QY	1 TACTGGAAGGTGGCGCTCTCTGCGTGTATACATGACATCCCACTGGCCCTGTGT	60
DB	12 TACTGGAAGGTGGCGCTCTCTGCGTGTATACATGACATCCCACTGGCCCTGTGT	71
QY	61 CTCGCTGCGCTCTCTGATACACAGCCTTCCGTGTATGTGAGGCGCCAGGCGTCAAGCGC	120

[illegible]

PN WO200198491-A2.
 XX 27-DEC-2001.
 XX 15-JUN-2001; 2001WO-EP06795.
 PF 19-JUN-2000; 2000EP-0112867.
 XX (HOFF) HOFEMANN LA ROCHE & CO AG F.
 PA (UYIN-) UNIV INSTELLING ANTWERPEN UIA.
 XX
 PI Bailemans W, Ebeling M, Foerzler D, Patel N, Van Hul W,
 PI Vickers BH;
 DR WPI: 2002-139789/18.
 DR P-PSDB; AAB17089.
 PT Novel genetic polymorphisms in the Van Buchem-sclerosteosis disease
 PT region that are associated with abnormal bone formation useful for
 PT diagnosis and assessment of osteoporosis or sclerosteosis in humans
 PS Claim 10; Fig 1B; 70pp; English.
 XX
 CC The invention relates to a nucleic acid encoding osteolevin region
 CC polymorphisms. The invention also relates to genetic polymorphisms in
 CC the Van Buchem-sclerosteosis disease region that are associated with
 CC disorders resulting in either net excess bone formation or insufficient
 CC bone formation in humans. Osteolevin DNA is useful for screening for
 CC osteolevin polymorphisms associated with abnormal bone formation in a
 CC subject and for the presence of a heritably linked form of abnormal bone
 CC formation in a subject. By determining the presence of a polymorphism in
 CC the osteolevin nucleic acid sequence obtained from the subject.
 CC Osteolevin protein is useful for treating diseases associated with
 CC abnormal bone formation, such as sclerosteosis, Van Buchem's disease and
 CC Paget's disease. Nucleic acids which encode genes in the osteolevin
 CC region or their modified forms can also be used to generate either
 CC transgenic animals or knockout animals which are useful in the screening
 CC and development of therapeutically useful reagents. Osteolevin proteins
 CC are useful in pharmacological characterisation of novel modulators of the
 CC activity of protein and protein complexes. Human osteolevin gene is
 CC located on chromosome 17. The present sequence is human osteolevin cDNA.
 XX
 SQ Sequence 2271 BP; 573 A; 623 C; 599 G; 476 T; 0 other;
 Query Match 96.8%; Score 735; DB 24; Length 2271;
 Best Local Similarity 100.0%; Pred. No. 1.9e-133;
 Matches 735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 25 CTGAGCTGATACCATGACAGCTCCAGCTGCTGTCTCTGCTGCTGTGACACACA 84
 DB 1 CTGAGCTGATACCATGACAGCTCCAGCTGCTGTCTCTGCTGCTGTGACACACA 60
 QY 85 GCCTTCCTGTAGTGAAGGGGCGAGGGGTGGAGAGCGTTCAAGATGATGCCAGGAATC 144
 DB 61 GCCTTCCTGTAGTGAAGGGGCGAGGGGTGGAGAGCGTTCAAGATGATGCCAGGAATC 120
 QY 145 ATCCCGAGCTCGGAGAGATACCCGAGCTCCACCGAGCTGAGAGAACAAAGACCATG 204
 DB 121 ATCCCGAGCTCGGAGAGATACCCGAGCTCCACCGAGCTGAGAGAACAAAGACCATG 180
 QY 205 AACCGGGGAGAAAGAGAGGGGAGCTCCCAACACCCCTTTGAGACAAAGAGTGTCC 264
 DB 181 AACCGGGGAGAAAGAGAGGGGAGCTCCCAACACCCCTTTGAGACAAAGAGTGTCC 240
 QY 265 GAGTACAGCTGCGGAGCTGCACTTCAACCGCTACGTAAGTGAAGTGGCCGCTGCGGAGC 324
 DB 241 GAGTACAGCTGCGGAGCTGCACTTCAACCGCTACGTAAGTGAAGTGGCCGCTGCGGAGC 300
 QY 325 GCCAAGCGGTCACGAGCTGTGTGCTCCGAGCAGTGGCGCCGCGGCGCTCTGCC 384
 DB 301 GCCAAGCGGTCACGAGCTGTGTGCTCCGAGCAGTGGCGCCGCGGCGCTCTGCC 360
 QY 385 AACGCCATCGGCGCGGCAAGTGTGGGACCTAGTGGGCCCGACCTTCGCTGCATCCCC 444

DB 361 AACGCCATCGGCGCGGCAAGTGTGGGACCTAGTGGGCCCGACTTCCGCTGCATCCCC 420
 QY 445 GACCGCTACCGGCGGAGCGGAGCTGCTGTGTCCCGGTGTGAAGGCCCGCGCGC 504
 DB 421 GACCGCTACCGGCGGAGCGGAGCTGCTGTGTCCCGGTGTGAAGGCCCGCGCGC 480
 QY 505 CGCAAGTGTGCGGCTGTGTGAGCTCGTGAAGTGAAGCGCCTACCGCTTCCACACACAG 564
 DB 481 CGCAAGTGTGCGGCTGTGTGAGCTCGTGAAGTGAAGCGCCTACCGCTTCCACACACAG 540
 QY 565 TCGAGCTCAAGAGATTTGGGAGACGAGGCCCTGCGCGGAGAAAGGGCCGGAACCCGCG 624
 DB 541 TCGAGCTCAAGAGATTTGGGAGACGAGGCCCTGCGCGGAGAAAGGGCCGGAACCCGCG 600
 QY 625 CCGCGGCGCGGAGCGCAAGGCCAGAGCGAGCTGAGAAAGCCTACTAGAGCCCG 684
 DB 601 CCGCGGCGCGGAGCGCAAGGCCAGAGCGAGCTGAGAAAGCCTACTAGAGCCCG 660
 QY 685 CCGGAGCCCTTCCCAAGCGGCGGCGCCCGGCTGAAACCGCGGCCCACTTTCTGTC 744
 DB 661 CCGGAGCCCTTCCCAAGCGGCGGCGCCCGGCTGAAACCGCGGCCCACTTTCTGTC 720
 QY 745 CTCTGCGCGTGTGTT 759
 DB 721 CTCTGCGCGTGTGTT 735
 RESULT 9
 AAA94051
 ID AAA94051 standard; DNA; 642 BP.
 XX AAA94051;
 AC
 XX
 DT 15-JAN-2001 (first entry)
 XX
 DE Human DAN/Cerberus-related protein 6 (hDCR6) cDNA exons 1 and 4.
 XX
 KW Human; DNA/Cerberus-related protein 6; hDCR6; morphogenic protein;
 KW antagonist; BMP; cell growth; cell differentiation; bone formation;
 KW gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..642
 FT /*tag= a
 FT /product= "hDCR6 #2"
 FT /partial
 XX
 XX WO200055193-A2.
 XX
 PD 21-SEP-2000.
 XX
 PF 02-MAR-2000; 2000WO-US05537.
 XX
 PR 12-MAR-1999; 99US-0124118.
 XX
 PA (REGG-) REGENERON PHARM INC.
 XX
 PI Economides AN;
 XX
 DR WPI: 2000-638179/61.
 DR P-PSDB; AAB26106.
 XX
 PT Novel isolated, human DNA/Cerberus related protein 6 which include
 PT natural homologue, and polypeptides comprising DCR6 domain and nucleic
 PT acids encoding the proteins which are useful as probes and primers -
 PS Claim 2; Fig 3; 40pp; English.
 XX
 CC The present sequence comprises exons 1 and 4 of the human
 CC DAN/Cerberus-related protein 6 (hDCR6) coding sequence. It was isolated

FT	/note= "This polymorphism results in an amino acid
FT	change from Arg to Xaa (stop codon)"
FT	replace (2190, T)
FT	/*tag= j
FT	/note= "This polymorphism results in an amino acid
FT	change from Trp to Xaa (stop codon)"
FT	replace (2408, G)
FT	/*tag= k
FT	variation
FT	replace (3539, T)
FT	/*tag= l
FT	variation
FT	replace (3944, T)
FT	/*tag= m
FT	variation
FT	replace (4425, G)
FT	/*tag= n
FT	variation
FT	replace (5097, A)
FT	/*tag= o
FT	complement (5099..2342)
FT	/*tag= p
FT	intron
FT	complement (5100..5331)
FT	/*tag= q
FT	exon
FT	/*tag= r
FT	number= 1
FT	replacement (5292, T)
FT	/*tag= x
FT	promoter
FT	complement (5366..5405)
FT	/*tag= s
FT	variation
FT	replace (5666, A)
FT	/*tag= t
FT	variation
FT	replace (5696, T)
FT	/*tag= u
FT	variation
FT	replace (5833, C)
FT	/*tag= v
FT	variation
FT	replace (6070, G)
FT	/*tag= w
FT	variation
FT	replace (6392, G)
FT	/*tag= x
FT	variation
FT	replace (6474, G)
FT	/*tag= y
FT	variation
FT	replace (6718..6719, CTCCT)
FT	/*tag= z
FT	variation
FT	replace (6737, A)
FT	/*tag= aa
FT	variation
FT	replace (6867, C)
FT	/*tag= ab
FT	variation
FT	replace (6926, G)
FT	/*tag= ac
FT	variation
FT	replace (6927, T)
FT	/*tag= ad
XX	
XX	WO200198491-A2.
XX	
XX	27-DEC-2001.
XX	
XX	15-JUN-2001; 2001WO-EP06795.
XX	
XX	19-JUN-2000; 2000EP-0112867.
XX	
XX	(HOFF) HOFFMANN LA ROCHE & CO AG F.
PA	(UYIN-) UNIV INSTELLING ANTWERPEN UIA.
XX	
PI	Balemans W, Ebeling M, Foernzler D, Patel N, Van Hul W,
P1	Vickey BH;
XX	
DR	WPI; 2002-139789/18.
DR	
XX	P-PSDB; AAEL7089.
CC	
CC	Novel genetic polymorphisms in the Van Buchem-sclerosteosis disease
PT	region that are associated with abnormal bone formation useful for
PT	diagnosis and assessment of osteoporosis or sclerosteosis in humans
XX	
PS	Claim 11; Fig 3; 70pp; English.
XX	
XX	The invention relates to a nucleic acid encoding osteolevin region
CC	polymorphisms. The invention also relates to genetic polymorphisms in

the Van Buchem-sclerosteosis disease region that are associated with disorders resulting in either net excess bone formation or insufficient bone formation in humans. Osteolevin DNA is useful for screening for osteolevin polymorphisms associated with abnormal bone formation in a subject and for the presence of a heritably linked form of abnormal bone formation in a subject, by determining the presence of a polymorphism in the osteolevin nucleic acid sequence obtained from the subject. Osteolevin protein is useful for treating diseases associated with abnormal bone formation, such as sclerosteosis, Van Buchem's disease and Paget's disease. Nucleic acids which encode genes in the osteolevin region or their modified forms can also be used to generate either transgenic animals or knockout animals which are useful in the screening and development of therapeutically useful reagents. Osteolevin proteins are useful in pharmacological characterisation of novel modulators of the activity of protein and protein complexes. Human osteolevin gene is located on chromosome 17. The present sequence is human osteolevin gene.

Query Match	68.1%;	Score 516.8;	DB 24;	Length 7099;
Best Local Similarity	96.0%;	Pred. No. 2.9e-91;		
Matches 530;	Conservative 0;	Mismatches 22;	Indels 0;	Gaps 0;

QY	208	CGGGGAGAAACGAGAGGGGGGCGCTTCCCAACAACCCCTTTAGACAACAAGAGTGTCCAG	267
Db	2330	CAGGGGTGGCCAGGGGGGGCGCACCTTCACGCGGGCTCTTCTCCACAGAGTGTCCAG	2331
QY	268	TACAGCTGCGGAGAGCTSCACTTCAACCCTGCTACGTAACCAATGGAGCGTGCAGAGCGC	327
Db	2330	TACAGCTGCGGAGAGCTSCACTTCAACCCGCTACGTAACCAATGGAGCGTGCAGAGCGC	2271
QY	328	AAGCCGCTCACCGAGCTGTGTCTCCGGCCAGTGGGCGCGCGCGCTGTGCTCCCAAC	387
Db	2270	AAGCCGCTCACCGAGCTGTGTCTCCGGCCAGTGGGCGCGCGCGCTGTGCTCCCAAC	2211
QY	388	GGCATTCGGCGCGGCAATGTGTGGAGACTTACTGTGGGCGCCGACTTCCGTGATCCCGAC	447
Db	2210	GGCATTCGGCGCGGCAATGTGTGGAGACTTACTGTGGGCGCCGACTTCCGTGATCCCGAC	2151
QY	448	CGCTACCGGCGGACGCGGTGAGCTGTGTATCCCGGTGATGAGCGCCGCGCGCGC	507
Db	2150	CGCTACCGGCGGACGCGGTGAGCTGTGTATCCCGGTGATGAGCGCCGCGCGCGC	2091
QY	508	AAGTGTGCGCTGTGTGCTTCGTGCAATGTCAGACGCGCTCACCCGCTTCCACAACCAATCG	567
Db	2090	AAGTGTGCGCTGTGTGCTTCGTGCAATGTCAGACGCGCTCACCCGCTTCCACAACCAATCG	2031
QY	568	GAGCTCAAGAGACTTTGGGACCGAGGCGCTCGGCGCGCAAGAGGCGCGAAGCCGCGGCGC	627
Db	2030	GAGCTCAAGAGACTTTGGGACCGAGGCGCTCGGCGCGCAAGAGGCGCGAAGCCGCGGCGC	1971
QY	628	CGCGCCCGGAGGCGCAAAAGCCAACAGGCGGAGCTGTGAGAAAGCTTACTAGAGCCCGCC	687
Db	1970	CGCGCCCGGAGGCGCAAAAGCCAACAGGCGGAGCTGTGAGAAAGCTTACTAGAGCCCGCC	1911
QY	688	GGCGCCCTTCCCAACCGGCGGCGCGCGGCGCTGAAACCGCGCGCCACATTTCTGTCTC	747
Db	1910	GGCGCCCTTCCCAACCGGCGGCGCGCGGCGCTGAAACCGCGCGCCACATTTCTGTCTC	1851
QY	748	TGCGGCTGTGTT 759	
Db	1850	TGCGGCTGTGTT 1839	

```

RESULT 14
AAA29064
ID      AAA29064 standard; DNA; 9301 BP.
XX
XX      AAA29064;
XX
XX      12-SEP-2000 (first entry)
DT
DE      Human TGF-beta binding protein (BEER) genomic DNA.

```

XX osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KW BEER; chromosome 17q12-21; gene therapy; antisense therapy; fracture;
 XX bone mineralization; ds.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT exon 162..428
 FT /*tag= a
 FT /number= 1
 FT CDS 209..3608
 FT /*tag= b
 FT /product= TGF-beta_binding_protein
 FT intron 429..3185
 FT /*tag= c
 FT exon 3186..5219
 FT /*tag= d
 FT /number= 2
 PN MO20032773-A1.
 XX 08-JUN-2000.
 XX 24-NOV-1999; 99MO-US27990.
 XX 27-NOV-1998; 98US-0110283.
 PA (DARW-) DARWIN DISCOVERY LTD.
 XX Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepel BW;
 PI Van Ness J, Winkler DG;
 DR WPI; 2000-412321/35.
 XX P-PSDB; AAY96429.
 PT Nucleic acids (1) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures
 PS Disclosure; Page 87-96; 162pp; English.
 XX This DNA encodes a human transforming growth factor-beta (TGF-beta)
 CC binding protein designated BEER. The hBEER gene has been localized
 CC to the chromosome 17q12-21. The cDNA and protein may be used for
 CC prevention, treatment and diagnosis of diseases associated with
 CC inappropriate BEER expression. For example, they may be used to treat
 CC disorders associated with decreased TGF-beta BP expression. The cDNA or
 CC vectors may be administered to treat diseases by rectifying mutations or
 CC deletions in a patient's genome that affect the activity of BEER by
 CC expressing inactive proteins or to supplement the patients own production
 CC of BEER polypeptides. The nucleic acids may be used for recombinant
 CC production of BEER, gene therapy, antisense therapy, as probes for
 CC diagnostic assays and for functional studies. BEER may be used to raise
 CC antibodies and for identification of BEER modulators. BEER antagonists
 CC may be used to increase bone mineral content for the treatment of
 CC disorders such as osteopenia, osteoporosis, fractures and other
 CC disorders associated with low mineral content.
 SQ Sequence 9301 BP; 2276 A; 2479 C; 2503 G; 2043 T; 0 other;

Query Match 68.1%; Score 516.8; DB 21; Length 9301;
 Best Local Similarity 96.0%; Pred. No. 2.9e-91;
 Matches 530; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 208 CGGGCGGAGAAAGGAGGGGGCTCCACACACCCCTTTGAACCAAGAGCTGCCAG 267
 DB 3138 CAGGGGTGGCCAGGGGGGCGCACCCCTCCTCCACAGACCTGTCCGAG 3197
 QY 268 TACAGTGGCGAGAGCTTGCATCCGCTAGTACCGATGGGCGGCGCGAGCGCC 327
 DB 3198 TACAGTGGCGAGAGCTTGCATCCGCTAGTACCGATGGGCGGCGCGAGCGCC 3257

QY 328 AAGCGGTACCGAGCTGTGTGCTCCGGCCAGTGCAGCCCGGCGCTGTGCCCAAC 387
 DB 3258 AAGCGGTACCGAGCTGTGTGCTCCGGCCAGTGCAGCCCGGCGCTGTGCCCAAC 3317
 QY 388 GCCATGGCCCGGCGCAAGTGTGGGCACTAGTGGGCCCCCACTTCCGCTGATCCCGGAC 447
 DB 3318 GCCATGGCCCGGCGCAAGTGTGGGCACTAGTGGGCCCCCACTTCCGCTGATCCCGGAC 3377
 QY 448 CGCTACCGCGGCGCAAGCTGTGTGCTCCGGTGGTGAAGGCGCGCGCGCG 507
 DB 3378 CGCTACCGCGGCGCAAGCTGTGTGCTCCGGTGGTGAAGGCGCGCGCGCGCG 3437
 QY 508 AAGTGTGCGCTGTGTGTGCTGTGCAAGTGCAGAGCGCTTCCACCAACAGTGG 567
 DB 3438 AAGTGTGCGCTGTGTGTGCTGTGCAAGTGCAGAGCGCTTCCACCAACAGTGG 3497
 QY 568 GAGCTCAAGGACTTTCGGGACCGGAGCGCTTGGCCCGCAAGAGGCGCGAGCC 627
 DB 3498 GAGCTCAAGGACTTTCGGGACCGGAGCGCTTGGCCCGCAAGAGGCGCGAGCC 3557
 QY 628 CGCGCCGCGAGCGCAAGCGCAAGCGAGCGAGCTGGAGAAAGCGCTTAGAGCCGCGCC 687
 DB 3558 CGCGCCGCGAGCGCAAGCGAGCGAGCGAGCTGGAGAAAGCGCTTAGAGCCGCGCC 3617
 QY 688 GCGCCCTCTCCACCGCGGCGCGCGCGCGCTTGAACCGCGCGCCCAATTCTGTCTC 747
 DB 3618 GCGCCCTCTCCACCGCGGCGCGCGCGCGCTTGAACCGCGCGCCCAATTCTGTCTC 3677
 QY 748 TCGCGTGTGTTT 759
 DB 3678 TCGCGTGTGTTT 3689

RESULT 15
 ID AAA29058
 AC AAA29058 standard; cDNA; 638 BP.
 AC AAA29058;
 DT 12-SEP-2000 (first entry)
 DE Murine TGF-beta binding protein (BEER) cDNA.
 XX osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
 KW BEER; gene therapy; antisense therapy; fracture; bone mineralization; se.
 OS Mus musculus.
 FH Key Location/Qualifiers
 FT CDS 1..636
 FT /*tag= a
 FT /product= TGF-beta_binding_protein
 PN MO20032773-A1.
 XX 08-JUN-2000.
 XX 24-NOV-1999; 99MO-US27990.
 XX 27-NOV-1998; 98US-0110283.
 PA (DARW-) DARWIN DISCOVERY LTD.
 XX Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepel BW;
 PI Van Ness J, Winkler DG;
 DR WPI; 2000-412321/35.
 DR P-PSDB; AAY96432.
 PT Nucleic acids (1) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures
 XX

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 28, 2003, 18:49:00 ; Search time 90.8624 Seconds
(without alignments)
7108.427 Million cell updates/sec

Title: US-09-867-274-1

Perfect score: 759
Sequence: 1 tactgaagtcgctgccc.....ctgcctctgctgctgtt 759

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 574371 seqs, 425486471 residues

Total number of hits satisfying chosen parameters: 1148742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	759	100.0	759	US-09-867-274-1	Sequence 1, Appl1
2	501.6	66.1	636	US-09-867-274-3	Sequence 3, Appl1
3	422	55.6	422	US-09-864-761-30988	Sequence 30988, A
4	355	46.8	392	US-09-864-761-14440	Sequence 14440, A
5	82.6	10.9	954	US-10-152-661-430	Sequence 430, App
6	82.6	10.9	954	US-09-866-050A-430	Sequence 430, App
7	82.6	10.9	962	US-10-152-661-40	Sequence 40, Appl
8	82.6	10.9	962	US-10-152-661-209	Sequence 209, App
9	82.6	10.9	962	US-09-866-050A-40	Sequence 40, Appl
10	82.6	10.9	962	US-09-866-050A-209	Sequence 209, App
11	69.6	9.2	900	US-09-853-625B-1	Sequence 1, Appl1
12	69.6	9.2	1592	US-09-745-288-92	Sequence 92, Appl1
13	52	6.9	4020	US-09-796-679-5	Sequence 5, Appl1
14	49.6	6.5	4257	US-09-825-288A-1	Sequence 1, Appl1
15	49	6.5	4098	US-09-962-436-37	Sequence 37, Appl
16	48.6	6.4	174424	US-09-967-768A-314	Sequence 314, App
17	47.4	6.2	2307	US-09-893-519A-87	Sequence 87, Appl
18	47	6.2	1514	US-09-976-740-45	Sequence 45, Appl
19	47	6.2	1514	US-10-023-529-45	Sequence 45, Appl

20	47	6.2	1614	US-10-023-523-45	Sequence 45, Appl
21	47	6.2	12425	US-09-976-740-50	Sequence 50, Appl
22	47	6.2	12425	US-10-023-529-50	Sequence 50, Appl
23	47	6.2	12425	US-10-023-523-50	Sequence 50, Appl
24	46.8	6.2	1910	US-09-789-836-1	Sequence 1, Appl1
25	46.8	6.2	2198	US-10-044-205A-1	Sequence 1, Appl1
26	46.8	6.2	2249	US-10-217-745-5	Sequence 5, Appl1
27	46.8	6.2	3651	US-09-964-469-3	Sequence 3, Appl1
28	44	5.8	14800	US-09-954-456-1601	Sequence 1601, Ap
29	44	5.8	88421	US-09-976-059-1	Sequence 1, Appl1
30	43.8	5.8	1275	US-09-815-242-4042	Sequence 4042, Ap
31	43.6	5.7	440	US-10-184-644-202	Sequence 202, App
32	43.6	5.7	8459	US-09-817-913-8	Sequence 8, Appl1
33	43.6	5.7	8459	US-09-817-538-8	Sequence 8, Appl1
34	43.4	5.7	8867	US-09-932-367A-22	Sequence 22, Appl
35	43.4	5.7	88421	US-09-976-059-1	Sequence 1, Appl1
36	43.2	5.7	13029	US-09-815-242-4052	Sequence 4052, Ap
37	43.2	5.7	33023	US-09-880-107-3350	Sequence 30, Appl
38	42.8	5.6	13842	US-09-860-846-30	Sequence 30, Appl
39	42.8	5.6	13842	US-09-861-289-30	Sequence 30, Appl
40	42.8	5.6	36778	US-09-860-846-5	Sequence 5, Appl1
41	42.8	5.6	36778	US-09-861-289-5	Sequence 5, Appl1
42	42.2	5.6	5173	US-09-880-107-3356	Sequence 3356, Ap
43	42.2	5.6	12606	US-09-957-974-2	Sequence 2, Appl1
44	42.2	5.6	17862	US-10-092-154-1313	Sequence 1313, Ap
45	42.2	5.6	17862	US-09-764-847-1313	Sequence 1313, Ap

ALIGNMENTS

RESULT 1
US-09-867-274-1
Sequence 1, Application US/09867274
Patent No. US20020106650A1
GENERAL INFORMATION:
APPLICANT: Paszty, Christopher
TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
FILE REFERENCE: 0101737428
CURRENT APPLICATION NUMBER: US/09/867,274
CURRENT FILING DATE: 2001-05-29
PRIOR APPLICATION NUMBER: US 60/208,550
PRIOR FILING DATE: 2000-06-01
PRIOR APPLICATION NUMBER: US 60/223,542
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 759
TYPE: DNA
ORGANISM: Homo sapiens
US-09-867-274-1

Query Match	100.0%;	Score 759;	DB 10;	Length 759;
Best Local Similarity	100.0%;	Pred. No. 1.2e-184;		
Matches 759;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	TACTGGAAGTGGGCTGCTCTCTGCTGTAACATGACGCTCCACATGGCCCTGTGT	60	
DB	1	TACTGGAAGTGGGCTGCTCTCTGCTGTAACATGACGCTCCACATGGCCCTGTGT	60	
QY	61	CTCGCTCTCGCTGCTGTAACATGACGCTCCACATGGCCCTGTGT	120	
DB	61	CTCGCTCTCGCTGCTGTAACATGACGCTCCACATGGCCCTGTGT	120	
QY	121	TTCAAGAAATGATGCTGCAAGAAATCATCCCTGAGCTCGAGATGACCCGAGCTTCAACG	180	
DB	121	TTCAAGAAATGATGCTGCAAGAAATCATCCCTGAGCTCGAGATGACCCGAGCTTCAACG	180	
QY	181	GAGCTGGAACCAACCAAGATGACATGACGCGGAGGAGGAGGCGGCTCCACACAC	240	
DB	181	GAGCTGGAACCAACCAAGATGACATGACGCGGAGGAGGAGGCGGCTCCACACAC	240	

QY 241 CCCTTTGAGACCAAGAGCTGTCCGAGTACAGTCCCGGAGCTGACCTTACACCCCGTAC 300
|
|
|
Db 241 CCCCTTTAGAACAAAGAGCTGTCCGAGTACAGTCCCGGAGCTGACCTTACACCCCGTAC 300
|
|
|
QY 301 GTGACCGATGGGCGGTCCGAGCGCCGCAAGCGGTGTCAACGAGTGTGTCTCCGCGCAG 360
|
|
|
Db 301 GTGACCGATGGGCGGTCCGAGCGCCGCAAGCGGTGTCAACGAGTGTGTCTCCGCGCAG 360
|
|
|
QY 361 TGGCGCCCGGCGCGCTGTGCTCCCAACGCAATGCGCGCGCAAGTGTGTGCACTAGT 420
|
|
|
Db 361 TGGCGCCCGGCGCGCTGTGCTCCCAACGCAATGCGCGCGCAAGTGTGTGCACTAGT 420
|
|
|
QY 421 GGGGCGGACCTTCCGCTGATCCCGACCGCTACCGCGCGGAGCGGTGACGCTGTGT 480
|
|
|
Db 421 GGGGCGGACCTTCCGCTGATCCCGACCGCTACCGCGCGGAGCGGTGACGCTGTGT 480
|
|
|
QY 481 CCCGATGATGAGCGCGCGCGCGCAAGAGTGGCTGTGTGCTGTGCAAGTGCAG 540
|
|
|
Db 481 CCCGATGATGAGCGCGCGCGCGCAAGAGTGGCTGTGTGCTGTGCAAGTGCAG 540
|
|
|
QY 541 CGCCTCACCCGCTTCCACACCAAGTCCGAGCTCAAGACTTCCGGAACCGAGCCGCTCGG 600
|
|
|
Db 541 CGCCTCACCCGCTTCCACACCAAGTCCGAGCTCAAGACTTCCGGAACCGAGCCGCTCGG 600
|
|
|
QY 601 CCGGAGAAAGGCGCGGAGCGCGCGCGCGCGCGCGGAGCGCAAGCGCAACGAGCCGAG 660
|
|
|
Db 601 CCGGAGAAAGGCGCGGAGCGCGCGCGCGCGCGCGGAGCGCAAGCGCAACGAGCCGAG 660
|
|
|
QY 661 CTGAGAGAACGCTTACTAGAGCCCGCGCGCTTCCCAACCGGCGGCGCGCGCGCT 720
|
|
|
Db 661 CTGAGAGAACGCTTACTAGAGCCCGCGCGCTTCCCAACCGGCGGCGCGCGCGCT 720
|
|
|
QY 721 GAACCCCGCGCGCGCAATTTCTGTCTCTGCGCGTGTGT 759
|
|
|
Db 721 GAACCCCGCGCGCGCAATTTCTGTCTCTGCGCGTGTGT 759
|
|
|

RESULT 2

US-09-867-274-3
; Sequence 3, Application US/09867274
; Patent No. US20020106650A1
; GENERAL INFORMATION:
; APPLICANT: Pastry, Christopher
; APPLICANT: Geo, Xongming
; TITLE OF INVENTION: Cysteine Knot Polypeptides: Cloaked-2 Molecules and Uses Thereof
; FILE REFERENCE: 01017/37428
; CURRENT APPLICATION NUMBER: US/09/867,274
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/208,550
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: US 60/223,542
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 636
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-867-274-3

Query March 66.1%; Score 501.6; DB 10; Length 636;
Best Local Similarity 87.5%; Pred. No. 4.3e-119;
Matches 562; Conservative 0; Mismatches 74; Indels 6; Gaps 1;

QY 37 ATGCAAGTCCCACTGAGCCCTGTGTCTGTCTGTCTGTGTATACAGAGCTTCCGTGTA 96
|
|
|
Db 1 ATGCAAGTCCCACTGAGCCCTGTGTCTGTCTGTGTATACAGAGCTTCCGTGTA 96
|
|
|
QY 97 GTGAGGCGCGAGGCGTGCAGAGCTTCAAGATATGCGACGAAATATATCCCGAGCTC 156
|
|
|
Db 61 GTGAGGCGCGAGGCGTGCAGAGCTTCAAGATATGCGACGAAATATATCCCGAGCTC 120
|
|
|

QY 157 GGAGAGTACCCCGAGCTCTCCACCGAGCTGGAGAACAAAGACCATGAACCCGCGGAG 216
|
|
|
Db 121 GGAGAGTACCCCGAGCTCTCTCC-----TGAGAACCAACGAGCATGAACCGGCGGAG 174
|
|
|
QY 217 AACGAGGCGCGGCTCTCCACCAACCCCTTTGAGACCAAGAGCTGTCCGAGTACGCTGC 276
|
|
|
Db 175 AATGAGGAGAGACTCTCCCAACCATCTATGACCGCAAAAGATGTGTCGAGTACGCTGC 234
|
|
|
QY 277 CGCAGAGTGCACCTTACCCCGTACTGTACCGATGGGCGGTGCGGCGGAGCGGCGGTC 336
|
|
|
Db 235 CGCAGAGTGCACCTTACCCCGTACTGTACCGATGGGCGGTGCGGCGGAGCGGCGGTC 294
|
|
|
QY 337 ACCGAGCTGTGTGTCTCCGCGCAGTGCAGCGCGCGCGCTGTGCGCAAGCCATCGGC 396
|
|
|
Db 295 ACCGAGTGTGTGTCTCCGCGCAGTGCAGCGCGCGCGCTGTGCGCAAGCCATCGGC 354
|
|
|
QY 397 CGCGCAAGTGTGTGCGACCTTATGTGGCGCGCACTTCCGCTGTATCCCGACCGCTACCGC 456
|
|
|
Db 355 CGCGTGAAGTGTGTGCGCGCGCGCAACGAGCCGATTTCCGCTGTATCCCGATCGCTACCGC 414
|
|
|
QY 457 GGGCAGCGCGTGCAGCTGTGTGTCCGCTGTGTGAGGCGCGCGCGCGCAAGTGTGC 516
|
|
|
Db 415 GGGCAGCGCGTGCAGCTGTGTGTCCGCGCGCGCGCGCGCGCGCTGTGCGCAAGTGTGC 474
|
|
|
QY 517 CTGTGTGCTCTGTGCAAGTGAAGCGCTTCAACCGCTTCCCAACCAAGTGCAGCTCAAG 576
|
|
|
Db 475 CTGTGTGCTCTGTGCAAGTGAAGCGCTTCAACCGCTTCCCAACCAAGTGCAGCTCAAG 534
|
|
|
QY 577 GACTTGGGAGCGGAGCGGCTGTGCGCGCGCAAGAGGCGCGGCGCGCGCGCGCGCG 636
|
|
|
Db 535 GACTTGGGAGCGGAGCGGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 594
|
|
|
QY 637 AGCGCAAGCAACGAGCGGAGCTGGAGAACGCTTACTAG 678
|
|
|
Db 595 GAGGCAAGCAACGAGCGGAGCTGGAGAACGCTTACTAG 636
|
|
|

RESULT 3

US-09-864-761-30988/c
; Sequence 30988, Application US/09864761
; Patent No. US2002048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wenheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30

```

; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 30988
; LENGTH: 422
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC003098.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.62
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.49
; OTHER INFORMATION: EST HUMAN HIT: BE613498.1, EVALUE 9.90e-02
; OTHER INFORMATION: SWISSPROT HIT: P45646, EVALUE 4.90e-01
; OTHER INFORMATION: NT HIT: AF074705.1, EVALUE 1.00e+00
; US-09-864-761-30988

```

```

Query Match      55.6%; Score 422; DB 10; Length 422;
Best Local Similarity 100.0%; Pred. No. 8.3e-99;
Matches 422; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 257 ACCTGTCCTCCAGTACAGTCCGCGGAGCTGACCTTCAACCCGCTACCGATGAGGCGCT 316
DB 422 ACCTGTCCTCCAGTACAGTCCGCGGAGCTGACCTTCAACCCGCTACCGATGAGGCGCT 363
QY 317 GCCGAGCGGCAAGCGGATGACGAGTGTGTCTCCGCGAGTGGCGCGCGCGCGC 376
DB 362 GCCGAGCGGCAAGCGGATGACGAGTGTGTCTCCGCGAGTGGCGCGCGCGCGC 303
QY 377 TGTGTCCTCCAGTACAGTCCGCGGAGCTGACCTTCAACCCGCTACCGATGAGGCGCT 436
DB 302 TGTGTCCTCCAGTACAGTCCGCGGAGCTGACCTTCAACCCGCTACCGATGAGGCGCT 243
QY 437 GCATCCCTCCAGTACAGTCCGCGGAGCTGACCTTCAACCCGCTACCGATGAGGCGCT 496
DB 242 GCATCCCTCCAGTACAGTCCGCGGAGCTGACCTTCAACCCGCTACCGATGAGGCGCT 183
QY 497 CGCGCGGCGCAAGTGTGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCT 556
DB 182 CGCGCGGCGCAAGTGTGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCT 123
QY 557 ACAACCAAGTGTGAGTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCT 616
DB 122 ACAACCAAGTGTGAGTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCTGTGAGGCT 63
QY 617 AGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 676
DB 62 AGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3
QY 677 AG 678
DB 2 AG 1

```

```

RESULT 4
US-09-864-761-14440/c
; Sequence 14440, Application us/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.

```

```

; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 14440
; LENGTH: 392
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC003098.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.62
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.49
; US-09-864-761-14440

```

```

Query Match      46.8%; Score 355; DB 10; Length 392;
Best Local Similarity 100.0%; Pred. No. 9.6e-82;
Matches 355; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 405 GTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 464
DB 392 GTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 333
QY 465 CGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 524
DB 332 CGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 273
QY 525 CTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 584
DB 272 CTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 213

```

QY 585 GACGAGGCGCTGCGCGCGAGAGGCGGAGCCCGCGCGCGAGAGCCCA 644
| | | | |
DB 212 GACGAGGCGCTGCGCGCGAGAGGCGGAGCCCGCGCGCGAGAGCCCA 153
| | | | |
QY 645 AGCAACGAGCGCGAGCTGAGAAAGCCTTACTAGAGCCCGCGCGCGCTCCCA 704
| | | | |
DB 152 AGCAACGAGCGCGAGCTGAGAAAGCCTTACTAGAGCCCGCGCGCGCTCCCA 93
| | | | |
QY 705 CGGCG 759
| | | | |
DB 92 CGGCG 38
| | | | |

RESULT 5

US-10-152-661-430
; Sequence 430, Application US/10152661
; Publication No. US20030022835A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c5
; CURRENT APPLICATION NUMBER: US/10/152,661
; PRIOR FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: 09/866,050
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/221,232
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: 60/206,650
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: 09/312,283
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: PCT/NZ99/00051
; PRIOR FILING DATE: 1999-04-29
; PRIOR APPLICATION NUMBER: 09/188,930
; PRIOR FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: 09/069,726
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 430
; LENGTH: 954
; TYPE: DNA
; ORGANISM: Mouse
US-10-152-661-430

Query Match 10.9%; Score 82.6; DB 9; Length 954;
Best Local Similarity 52.5%; Pred. No. 2.2e-12;
Matches 287; Conservative 0; Mismatches 239; Indels 21; Gaps 4;
QY 112 TGGCAGGCGCTTCAAGATGATGTCACGGAATCATCCCGAGCTCGAGAGTACCCCGAG 171
| | | | |
DB 99 TGTTCGCTTTTAAATGATGTCACGGAATCCTTATTACATGCTGTTAAACCTGTC 158
| | | | |
QY 172 CCTCCACGAGCTGAGAACCAAGACATGAACCGGCGGAGAACGAGGCGGCGCT 231
| | | | |
DB 159 CGGCGAC---ACCCAGAGCAAGACGACCTGATCAAGCAAGAAATGAGGAGGAGCAT 215
| | | | |
QY 232 CCCACACACCC---CTTGAAGACCAAGACGTCGAGTCCAGCTCCCGAGAGCTGCAC 288
| | | | |
DB 216 TTCAAGTACCTGAGCTGATCGAAACAGTCAGTTCAGTGGGCTCGAGGAACTTCGCG 275
| | | | |
QY 289 TTCAACCGCTACGTGACCGATGAGGCGGTCGCGAGCGCCAGAGCGGCTCACCGAGCTG 348
| | | | |
DB 276 TCCACCAATATCATTTTGGAGCGGCAAGTCACAGCATCAGCCTCTGAAAGAGCTG 335
| | | | |
QY 349 TGTTCGCGCAGTGGCGCGCGCGCGCTGCTGCGCAACCGCATCGGCGCGG----- 402
| | | | |

DB 336 TGCAGGCGGAGTGTCTTCCCTCGCGGTCTTCCCACTGATCGAGAGGCTACGGA 395
| | | | |
QY 403 -----AGTGTGCGACACTGATGAGGCGCGACTTCGCTGATCCCGAGCGCTACGCG 456
| | | | |
DB 396 ACAAGTACTGAGGCGGAGAGGCTCTAGAGTGGCGGTGTGTCAACGAGAAAGCGCGC 455
| | | | |
QY 457 GCGAGCGGCTGACGCTGCTGTGTCGCGGTGAGGCGCGCGCGCGCGAGGTGCGC 516
| | | | |
DB 456 ACCAGAGGATCTCACTGATGATGTCAGAGCGG---CAGCAGCGGACCTCAAAATACCC 512
| | | | |
QY 517 GTGTGCGCTGTGCAAGTGCAGGCGCTCACCGGCTTCCAAACAGTGGAGCTCAG 576
| | | | |
DB 513 GTGTGTCAGCGCTGCAAGTGCAGAGGTCACCGCTCAGCAACAGATCCAGCACAAAC 572
| | | | |
QY 577 GACTTCGAGACGAGCGGCTCGCGCGGAGAGGCGGAGCGCGCGCGCGCGCGCG 636
| | | | |
DB 573 TTGAAAGCGTGTGCTGCGCAGCCAGCGCGCCGACACACAGAGAGGAGAGCGCAGC 632
| | | | |
QY 637 AGCGCCA 643
| | | | |
DB 633 AATCCA 639
| | | | |

RESULT 6

US-09-866-050A-430
; Sequence 430, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; PRIOR FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 430
; LENGTH: 954
; TYPE: DNA
; ORGANISM: Mouse
US-09-866-050A-430

Query Match 10.9%; Score 82.6; DB 9; Length 954;
Best Local Similarity 52.5%; Pred. No. 2.2e-12;
Matches 287; Conservative 0; Mismatches 239; Indels 21; Gaps 4;
QY 112 TGGCAGGCGCTTCAAGATGATGTCACGGAATCATCCCGAGCTCGAGAGTACCCCGAG 171
| | | | |
DB 99 TGTTCGCTTTTAAATGATGTCACGGAATCCTTATTACATGCTGTTAAACCTGTC 158
| | | | |
QY 172 CCTCCACGAGCTGAGAACCAAGACATGAACCGGCGGAGAACGAGGCGGCGCT 231
| | | | |
DB 159 CGGCGAC---ACCCAGAGCAAGACGACCTGATCAAGCAAGAAATGAGGAGGAGCAT 215
| | | | |
QY 232 CCCACACACCC---CTTGAAGACCAAGACGTCGAGTCCAGCTCCCGAGAGCTGCAC 288
| | | | |
DB 216 TTCAAGTACCTGAGCTGATCGAAACAGTCAGTTCAGTGGGCTCGAGGAACTTCGCG 275
| | | | |
QY 289 TTCAACCGCTACGTGACCGATGAGGCGGTCGCGAGCGCCAGAGCGGCTCACCGAGCTG 348
| | | | |
DB 276 TCCACCAATATCATTTTGGAGCGGCAAGTCACAGCATCAGCCTCTGAAAGAGCTG 335
| | | | |
QY 349 TGTTCGCGCAGTGGCGCGCGCGCGCTGCTGCGCAACCGCATCGGCGCGG----- 402
| | | | |
DB 336 TGCAGGCGGAGTGTGCTGCGCGGCTGCTGCGCAACCGCATCGGCGCGG----- 402
| | | | |
QY 403 -----AGTGTGCGACACTGATGAGGCGCGGCTTCCGCTGATCCCGAGCGCTACCGC 456
| | | | |

Db 396 ACAAGTACTGAGCCGAGGAGAGCTCTCAGAGAGTGGCGGTGTCTCAACAACAAGCGGC 455
Qy 457 GCGCAGCGCGTGCAGCTGTGTCTCCGAGTGTGAGCGCGCGCGCGCAAGGTGGC 516
Db 456 ACCCAGAGATCCAGCTGTGTCTCAGAGAGG---CAGCAGCGCGCACTTACAAATCACC 512
Qy 517 CTGTGCGCTGTGCAAGTGCAGAGCGCTTCCAGTCCAGCAACAGTGCAGTCAAG 576
Db 513 GTGTCAAGCGCGTGCAGAGTGTACAGCGCTTCAAGACAGAGTCCAGCGCAAC 572
Qy 577 GACTTGGAGCCAGCGCGCTCCGCGCAGAGAGCGCGCGCGCGCGCGCGCG 636
Db 573 TTGAAAGCGTGTGTCCAGCAAGCCCGCCAGCAGCAGAGAGCGAGAGAGCGCAGC 632
Qy 637 AGCGCCA 643
Db 633 AATCCA 639

RESULT 7

US-10-152-661-40
; Sequence 40, Application US/10152661
; Publication No. US20030022835A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011CS
; CURRENT APPLICATION NUMBER: US/10/152,661
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: 09/866,050
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/221,232
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: 60/206,650
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: 09/312,283
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: PCT/NZ99/00051
; PRIOR FILING DATE: 1999-04-29
; PRIOR APPLICATION NUMBER: 09/188,930
; PRIOR FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: 09/069,726
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 962
; TYPE: DNA
; ORGANISM: Mouse
US-10-152-661-40

Query Match 10.9%; Score 82.6; DB 9; Length 962;
Best Local Similarity 52.5%; Pred. No. 2.2e-12;
Matches 287; Conservative 0; Mismatches 239; Indels 21; Gaps 4;

Qy 112 TGGCAGCGCTTCAAGATGATGTCAGGAATCATCCCGAGCTCGAGAGTACCCGAG 171
Db 107 TGTGGCTTTTAAATATGATGTCAGGAATCTTTTATTCATATGTGTTAACTGTTC 166
Qy 172 CCTCCACCGAGTGTGAGAACAAAGACCATTAACCGGCGGAGAAAGAGGCGGCT 231
Db 167 CCGGAC---ACCCAGCAGCAACAGCAGCCCTGATCAAGCCAGATGAGGAGCGCAT 223
Qy 232 CCCACACACCC---CTTGGAGACCAAGAGTGTCCGAGTACAGCTCGCGCGAGCTGCAC 288
Db 224 TTCAATGACACTGAGTGTGTCGAACAATGTCAGATTCAATGAGGAGTGCAG 283

Qy 289 TTCACCCGCTACGTGACCGATGAGCGCGTGTCCGAGCGCCAAAGCCGCTGAGTGTG 348
Db 284 TTCACCAATATTCGTCGAGCGGACAGTGTACACAGCATACAGCCCTCTTAAAGAGCTGGT 343
Qy 349 TGTCTCCGCTCAGTGTGCGCGCGCGCGCTGTGCTGCTCCAAAGCATTCGCGCGGCGC----- 402
Db 344 TGGCGGCGAGGTGTTCCTCCCTGCGGTGTCTTCCCACTGAATCGAGAGAGGCTACGGA 403
Qy 403 -----AAGTGTGCGCACTTATGAGCGCGCACTTCCGCTGATCCCGACCGCTACCGC 456
Db 404 ACAAGTACTGAGCCGAGAGAGTCTTACAGAGTGGCGGTGTCTCAACAACAAGCGGC 463
Qy 457 GCGCAGCGCGTGCAGCTGTGTCTCCGAGTGTGAGAGCGCGCGCGCGCGCAAGTGGC 516
Db 464 ACCCAGAGATCCAGCTGTGTCTCAGAGAGG---CAGCAGCGCGCACTTACAAATCACC 520
Qy 517 CTGTGCGCTGTGCAAGTGCAGAGCGCTTCCAGCTTCCAAACAGTGCAGTCAAG 576
Db 521 GTGTCAAGCGCGTGCAGAGTGTACAGCGCTTCAAGCAACAAGTTCAGCCAGC 580
Qy 577 GACTTGGAGCCAGCGCGCTCCGCGCAGAGAGCGCGCGCGCGCGCGCGCG 636
Db 581 TTGAAAGCGTGTGTCCAGCAAGCCCGCCAGCAGCAGAGAGCGGAGAGAGCGCAGC 640
Qy 637 AGCGCCA 643
Db 641 AATCCA 647

RESULT 8

US-10-152-661-209
; Sequence 209, Application US/10152661
; Publication No. US20030022835A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011CS
; CURRENT APPLICATION NUMBER: US/10/152,661
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: 09/866,050
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: 60/221,232
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: 60/206,650
; PRIOR FILING DATE: 2000-05-24
; PRIOR APPLICATION NUMBER: 09/312,283
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: PCT/NZ99/00051
; PRIOR FILING DATE: 1999-04-29
; PRIOR APPLICATION NUMBER: 09/188,930
; PRIOR FILING DATE: 1998-11-09
; PRIOR APPLICATION NUMBER: 09/069,726
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 209
; LENGTH: 962
; TYPE: DNA
; ORGANISM: Mouse
US-10-152-661-209

Query Match 10.9%; Score 82.6; DB 9; Length 962;
Best Local Similarity 52.5%; Pred. No. 2.2e-12;
Matches 287; Conservative 0; Mismatches 239; Indels 21; Gaps 4;

Qy 112 TGGCAGCGCTTCAAGATGATGTCAGGAATCATCCCGAGCTCGAGAGTACCCGAG 171

Db 107 TGTTCGCTTTTAAAAATGATGTCACAGAAATCCCTTTATTCATGTTGTTAAAACTGTG 166
QY 172 COTCCACCGAGCTGGAGAACAAACAGACATGAACCGGGGGGAGAAACGAGGGGCGCT 231
Db 167 CCGGAC---ACCCAGACGACACACCTTGAATCAAGCAGAGATGGAGGAGGAT 223
QY 232 CCCACCAACC---CTTTAGAACCAAAACGTCGAGTCAAGCTGCGGAGCTGAC 288
Db 224 TTCAAGTACACTGGACCTGGATCGAAACAGTGAAGTGGGCTGAGGAACTGCGG 283
QY 289 TTCAACCCGCTACGTGACCGATGGGCGCTGCGGACGCGCAAGCGGTGACCGAGCTGTG 348
Db 284 TCACCAAAATCATTTTCGAGCGGCGACAGTCAACAGCATCAACCTCTGAAGAGACTGTG 343
QY 349 TGCTCCGCGACGTGGGCGCGCGCGCTGCTGCGCAACGCGCATCGGCGCGGCGG----- 402
Db 344 TGCGCGGCGAGTGTCTTCCCTGCGGCTGCTTCCCACTGGAATCGAGAGAGGCTTACGGA 403
QY 403 -----AAGTGGGCGACCTAGTGGGCGCGACCTTCCGCTGCATCCCGACCGCTACCGC 456
Db 404 ACAAAATCTGGAGCGGAGGAGCTCTGAGAGTGGCGGTGTCAACAGACAAAGACGCGC 463
QY 457 GCGCAGCGGCTGCAAGTCTGTGTCGCGTGTGAGGCGCGCGCGCGCAAGGTGCGC 516
Db 464 ACCAGAGGATTCAGCTGCAAGTGTCAAGACG---CAGCAGCGCGACCTACAAATACACC 520
QY 517 CTGTCGCTGCTGCAAGTGAAGGCGCTCAACCGCTTCCCAACCAAGTGGAGCTGAG 576
Db 521 GTGTGTCAGGCTGCAAGTGAAGGAGTACACCGCTGACGACCAAGTCCAGCACAC 580
QY 577 GACTTCGAGACCGAGCGCTGCGCGCGCAAGAGGCGCGGAGCGCGCGCGCGCG 636
Db 581 TTTGAAAGCTGTGCGCGACCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 640
QY 637 AGCGCCA 643
Db 641 AAATCCA 647

RESULT 9

US-09-866-050A-40
; Sequence 40, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 962
; TYPE: DNA
; ORGANISM: Mouse
US-09-866-050A-40

Query Match 10.9%; Score 82.6; DB 9; Length 962;
Best Local Similarity 52.5%; Pred. No. 2.2e-12;

Matches 287; Conservative 0; Mismatches 219; Indels 21; Gaps 4;

QY 112 TGCGAGGCGCTTCAAGATGATGCCAGGAATCATCCCGGAGCTCGAGAGTACCCCGAG 171
Db 107 TGTTCGCTTTTAAAAATGATGTCACAGAAATCCCTTTATTCATGTTGTTAAAACTGTG 166
QY 172 COTCCACCGAGCTGGAGAACAAACAGACATGAACCGGGGGGAGAAACGAGGGGCGCT 231

Db 167 CCGGAC---ACCCAGACGACACAGCACCTTGAATCAAGCAGAGATGAGGAGGAGAT 223
QY 232 CCCACCAACC---CTTTAGAACCAAAACGTCGAGTCAAGCTGCGGAGCTGAC 288
Db 224 TTCAAGTACACTGGAGCTGGATCGAAACAGTGAAGTGGGCTGAGGAACTGCGG 283
QY 289 TTCAACCCGCTACGTGACCGATGGGCGCTGCGGACGCGCAAGCGGTGACCGAGCTGTG 348
Db 284 TCACCAAAATCATTTTCGAGCGGCGACAGTCAACAGCATCAACCTCTGAAGAGCTGTG 343
QY 349 TGCTCCGCGACGTGGGCGCGCGCGCTGCTGCGCAACGCGCATCGGCGCGGCGG----- 402
Db 344 TGCGCGGCGAGTGTCTTCCCTGCGGCTGCTTCCCACTGGAATCGAGAGAGGCTTACGGA 403
QY 403 -----AAGTGGGCGACCTAGTGGGCGCGACCTTCCGCTGCATCCCGACCGCTACCGC 456
Db 404 ACAAAATCTGGAGCGGAGGAGCTCTGAGAGTGGCGGTGTCAACAGACAAAGACGCGC 463
QY 457 GCGCAGCGGCTGCAAGTCTGTGTCGCGTGTGAGGCGCGCGCGCGCAAGGTGCGC 516
Db 464 ACCAGAGGATTCAGCTGCAAGTGTCAAGACG---CAGCAGCGCGACCTACAAATACACC 520
QY 517 CTGTCGCTGCTGCAAGTGAAGGCGCTCAACCGCTTCCCAACCAAGTGGAGCTGAG 576
Db 521 GTGTGTCAGGCTGCAAGTGAAGGAGTACACCGCTGACGACCAAGTCCAGACCAAC 580
QY 577 GACTTCGAGACCGAGCGCTGCGCGCGCAAGAGGCGCGGAGCGCGCGCGCGCGCG 636
Db 581 TTTGAAAGCTGTGCGCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 640
QY 637 AGCGCCA 643
Db 641 AAATCCA 647

RESULT 10

US-09-866-050A-209
; Sequence 209, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 209
; LENGTH: 962
; TYPE: DNA
; ORGANISM: Mouse
US-09-866-050A-209

Query Match 10.9%; Score 82.6; DB 9; Length 962;
Best Local Similarity 52.5%; Pred. No. 2.2e-12;

Matches 287; Conservative 0; Mismatches 219; Indels 21; Gaps 4;

QY 112 TGCGAGGCGCTTCAAGATGATGCCAGGAATCATCCCGGAGCTCGAGAGTACCCCGAG 171
Db 107 TGTTCGCTTTTAAAAATGATGTCACAGAAATCCCTTTATTCATGTTGTTAAAACTGTG 166
QY 172 COTCCACCGAGCTGGAGAACAAACAGACATGAACCGGGGGGAGAAACGAGGGGCGCT 231
Db 167 CCGGAC---ACCCAGACGACACACCTTGAATCAAGCAGAGATGGAGGAGGAT 223
QY 232 CCCACCAACC---CTTTAGAACCAAAACGTCGAGTCAAGCTGCGGAGGAGTGCAC 288

SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 us-09-853-625B-1

Query Match 9.2%; Score 69.6; DB 10; Length 900;
 Best Local Similarity 50.0%; Pred. No. 4.6e-09;
 Matches 239; Conservative 0; Mismatches 224; Indels 15; Gaps 2;

Db 224 TTGAGTAGCACTGGAGTGTGATCCAAACATGTCAGTTCAGTGGGCTGACAGGAAGTCCGG 283
 Qy 289 TTGACCCGCTACGAGTCCGATGGGCGTGCAGAGCGCAAGCCGGTACCCGAGCTGGTG 348
 Db 284 TCACCAAAATACATTTTGGACGGCGCATGACACAGCATCAGCCCTCTGAAGAGAGTGGTG 343
 Qy 349 TGTCCCGCCAGTGGGCGCGCGCGCTGCTGCCAACCCATCCGACCGCGCGC----- 402
 Db 344 TGGCGGGGCAAGTGTGCTTGGCCCTGCGGTGTCTTCCAACTGATTCGAGAGAGGCTTACGGA 403
 Qy 403 -----AAGTGTGGCGCACTTGTGGGCGCGGACTTCCGCTCATCTCCCGACCGCTACCG 456
 Db 404 ACAAAATACAGAGCCGAGAGAGCTCTCAGAGTGGCGGTGTCTCAACAGAGAGCGCG 463
 Qy 457 GCGCAGCGCGTGAAGTGTGTGTCCCGGTGTGAGGCGCGCGCGCGCGCAAGTGGCG 516
 Db 464 ACCCAGAGATTCAGTGTGCTGACGTGTCAAGAGCG---CAGCAGCGCACCTTACAAATAC 520
 Qy 517 CTGTGGCTCTGTGCAAGTGCAGAGCGCTTCCGCTTCCAAACAGTGCAGTCAAG 576
 Db 521 GTGTGTCAGCGGTGCAAGTGCAGAGAGTACACCGTACGACAAAGATCAGCCACAG 580
 Qy 577 GACTTGGGACCAAGCGCGCTCGCGCGCAGAGAGGCGGAGCGCGCGCGCGCGCG 636
 Db 581 TTGAAAGCGTGTGCGCAGCAGCGCGCGCGCGCGCGCGCGAGAGAGAGCGCAG 640
 Qy 637 AGCGCCA 643
 Db 641 AATCCA 647

RESULT 11
 us-09-853-625B-1
 ; Sequence 1, Application US/09853625B
 ; Patent No. US20020049304A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Haelling, Gregg A. and Adams, Mark D.
 ; TITLE OF INVENTION: Human CCN-Like Growth Factor
 ; NUMBER OF SEQUENCES: 20
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
 ; CECCHI, STEWART & OLSTEIN
 ; STREET: 6 BECKER FARM ROAD
 ; CITY: ROSELAND
 ; STATE: NEW JERSEY
 ; COUNTRY: USA
 ; ZIP: 07068
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5 INCH DISKETTE
 ; COMPUTER: IBM PS/2
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: WORD PERFECT 5.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/853,625B
 ; FILING DATE: 14-May-2001
 ; CLASSIFICATION: <Unknown>
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 09/053,587
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: MULLINS, J.G.
 ; REGISTRATION NUMBER: 33,073
 ; REFERENCE/DOCKET NUMBER: 325800-442
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 201-994-1700
 ; TELEFAX: 201-994-1744
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 900 BASE PAIRS
 ; TYPE: NUCLEIC ACID
 ; STRANDEDNESS: SINGLE
 ; TOPOLOGY: LINEAR
 ; MOLECULE TYPE: cDNA

SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 us-09-853-625B-1

Query Match 9.2%; Score 69.6; DB 10; Length 900;
 Best Local Similarity 50.0%; Pred. No. 4.6e-09;
 Matches 239; Conservative 0; Mismatches 224; Indels 15; Gaps 2;

Db 110 GGTGACAGCGCTTCAAGATGATGACAGGAATGATCCCGAGCTGGAGATCCCG 169
 Qy 176 GCTGTTGGCTTTTAAATATGATGACAGGAATCTTTATTCATATGTGTTAACTTG 235
 Db 170 ACCCTTCAACCGAGCTGAGAGAACAAAGACATGAAACCGGCGGAGAGAGCGCGC 229
 Qy 236 TTCAGACACACCCAGACAGCAACAGCATGTTGAATCAAGCCAGAAATGAGAGCGCAT 295
 Db 230 CTCGCCACCAACCCCTTTTGAACCAAGACGTGTCCGATGACAGCTGCGGAGTCACT 289
 Qy 296 TCAGTAAACATGATGATGAGAACTCGGGTTCAAGTGGGTGCGGGAACGCGTT 355
 Db 290 TCACCCGCTACGTGACCGATGGGCGGTGCGGAGCGCGGAGCGGCTGATGAT 349
 Qy 356 CCAACCAATACATCTCTGATGAGCGCATGACACGATGATGAGCTGTTGAAGAGCTGTGT 415
 Db 350 GCTCCGCGCAAGTGCAGCGCGCGCGCTGTGCCCCAACCGCATGCGCGCGC----- 402
 Qy 416 GTGTGCGGAGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 475
 Db 403 -----AAGTGTGGCGCACTTGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCG 457
 Qy 476 CAAAGTACTGAGAGAGAGAGAGCTCCAGAGTGGCGGTGTCAATACAAACCCGCTA 535
 Db 458 GCGCAGCGGTGACGTGTGTCCTCCGGTGTGAGAGCGCGCGCGCGCGCAAGTGGCGCC 517
 Qy 536 CCCAAGATACATCAGTGTGACAGTGTCCAAAGTGG---CAGCAGCGCACCTTACAAATACAG 592
 Db 518 TGTGGCGCTCTGTGCAAGTGCAGAGCGCTTCCGCTTCCAAACAGTGCAGTCAAG 575
 Qy 593 TAGTACATGCTGCAAGTGCAGAGAGTACACCGGACAGCAACAGATGCAAGTCAAG 650

RESULT 12
 us-09-745-288-92/c
 ; Sequence 92, Application US/09745288
 ; Patent No. US20010018058A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Reed, Steven G.
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Dillon, David C.
 ; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
 ; FILE REFERENCE: 210121.446D1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/745,288
 ; FILING DATE: 2000-12-19
 ; NUMBER OF SEQ ID NOS: 101
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 92
 ; LENGTH: 1692
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; us-09-745-288-92

Query Match 9.2%; Score 69.6; DB 10; Length 1692;
 Best Local Similarity 50.0%; Pred. No. 4.5e-09;
 Matches 239; Conservative 0; Mismatches 224; Indels 15; Gaps 2;

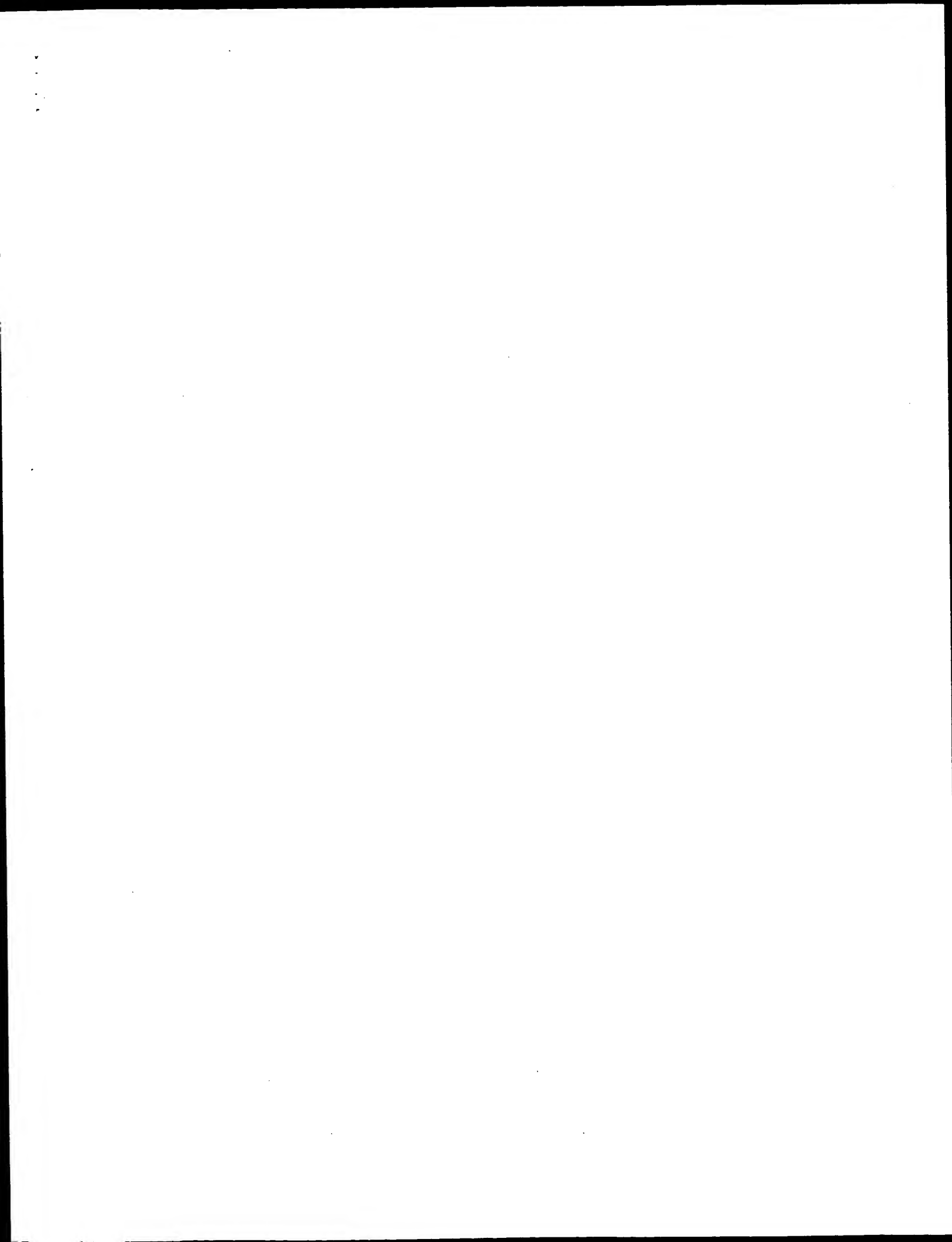
Db 110 GGTGACAGCGCTTCAAGATGATGACAGGAATGATCCCGAGCTGGAGATCCCG 169
 Qy 1582 GCTGTTGGCTTTTAAATATGATGACAGGAATCTTTATTCATATGTGTTAACTTG 1523
 Db 170 AGCTCCACCGAGCTGAGAGAACAAAGACATGAAACCGGCGGAGAGAGAGGCGCG 229
 Qy 1522 TTCCAGCACACCCAGACAGCAACGACGTGTTGAATCAAGCCAGAAATGAGAGCGCAT 1463

; PRIOR FILING DATE: 2000-09-25
 ; NUMBER OF SEQ ID NOS: 568
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 37
 ; LENGTH: 4098
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-962-436-37

Query Match 6.5%; Score 49; DB 10; Length 4098;
 Best Local Similarity 44.2%; Pred. No. 0.00077;
 Matches 202; Conservative 0; Mismatches 255; Indels 0; Gaps 0;

QY 268 TACAGCTGCGGAGCTCACTTACCCGCTTACGTGACCGATGGGCCGTGCCGAGCGCC 327
 Db 473 TCCGCGCCCGCCCGCCCGCCCGCCGAGCAGCGCCCGCCCGCTTCTCGGAAACCAAGTTC 414
 QY 328 AAGCGGTACCGAGCTGTGTGCTCCGCGCCAGTGCAGCCCGCGCGCTGTGCTCCCAAC 387
 Db 413 CACCTGGAAACGGCTCTGTGCTCGGGGTGGGCCCCAAGGGTCTTGCCAGCCGTCGCCGCG 354
 QY 388 GCCATCGGCGCGGCAAGTGTGCGACCTAGTGGCCCGACCTTCCGCTGCATCCCGAC 447
 Db 353 CGCCGCGGGGCGCTCATGCGGACCCCGCGCGCTCCCGCTCGCGGAGCGACATCTTC 294
 QY 448 CGCTACCGCGCGCAGCGGTGCACTGTGTCTCCGCTGTGAGGCGTCGCGCGCGCG 507
 Db 293 CGGCACCGAGGGCAAGCGCGGTGCGGCGAGTTCACCTGTGCTGCGGCAAGCGCGCG 234
 QY 508 AAGGTGCGCTGTGCTGTGCTGTGCAAGTGAAGCGCTCAACCGCTTCCACCAACGATCG 567
 Db 233 TGACGGGCTTCTCCCGGACCCCGGCAATCCCGGGGCGCCGAGAGAGGCGCTGTGGCG 174
 QY 568 GAGCTCAAGACTTTCGAGACCGAGCGCTCGGCGCGCAAGGGCGCGAGCCGCGCC 627
 Db 173 CGGCTTCATAGCTGCGGACCCCGCGGAACTGCAAGAGCCCGGAGACGTCTCGGCGAG 114
 QY 628 CGGCGCGGAGCGCGCAAGCGCAACCAAGCCGAGCTGAGAGACCTTACTAGAGCCCGCC 687
 Db 113 CAGGCGCTTCCACACCTTGGCGGCGGTGCGCACAGCGCGCAAGACAGCTGAGAGAGTCT 54
 QY 688 GCGCGCTTCCCGACCGGCGGCGCGCGCGCGCTGAAC 724
 Db 53 TTTCTCCCGCGCTGCGGGCGCGCTAGCTTGAC 17

Search completed: March 29, 2003, 01:33:21
 Job time : 98.8624 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 29, 2003, 00:44:05 ; Search time 2465.25 Seconds

(without alignments)
8960.144 Million cell updates/sec

Title: US-09-867-274-1

Perfect score: 759

Sequence: 1 tctggaaggtggtggtcc.....ctgtccctcgtggttt 759

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4139280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_stb:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_pl:*
25: em_pi:*
26: em_ro:*
27: em_stb:*
28: em_un:*
29: em_vi:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pin:*
35: em_hcg_rtd:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hcgo_hum:*
40: em_hcgo_mus:*
41: em_hcgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	759	100.0	759	6	AX323453 Sequence
2	759	100.0	2296	6	AX331844 Homo sapi
3	759	100.0	2323	9	AF326739 Homo sapi
4	757	99.7	2329	6	AX056687 Sequence
5	735	96.8	2271	6	AX342535 Sequence
6	505	66.5	7099	6	AX342537 Sequence
7	505	66.5	21501	9	AF326736 Homo sapi
8	505	66.5	80117	9	AC055813 Homo sapi
9	505	66.5	94752	9	AC003098 Homo sapi
10	220	29.0	17744	2	AC073954 Homo sapi
11	110	14.5	642	9	AF326742 Cercopit
12	83	10.9	532	4	AF326738 Bos tauru
13	81	10.7	93790	9	AF397423 Homo sapi
14	69	9.1	636	6	AX323455 Sequence
15	69	9.1	638	10	AF326740 Mus muscu
16	69	9.1	81806	2	AF326737 Mus muscu
17	69	9.1	110000	2	AC068782-2 Continuation (3 of
18	69	9.1	205277	2	AC012296 Mus muscu
19	69	9.1	208135	10	AC068807 Mus muscu
20	68	9.0	674	10	AF326741 Mus muscu
21	68	9.0	101804	2	AC098160 Rattus no
22	68	9.0	104898	2	AC121721 Rattus no
23	40	5.3	40	6	AX056701 Sequence
24	35	4.6	198508	10	AL591145 Mouse DNA
25	27	3.6	27	6	AX056700 Sequence
26	27	3.6	27	6	AX323473 Sequence
27	26	3.4	26	6	AX323460 Sequence
28	24	3.2	24	6	AX323459 Sequence
29	23	3.0	23	6	AX323474 Sequence
30	23	3.0	23	6	AX342540 Sequence
31	22	2.9	51575	2	AC023810 Mus muscu
32	22	2.9	170472	2	AC126215 Rattus no
33	22	2.9	194746	2	AC123544 Lemur cat
34	21	2.8	21	6	AX056699 Sequence
35	21	2.8	41	6	AX323472 Sequence
36	21	2.8	1862	6	AK093966 Homo sapi
37	21	2.8	2266	6	AR073028 Sequence
38	21	2.8	2266	9	HIMELK1A M2525 Homo sapien
39	21	2.8	2695	9	AB016193 Homo sapi
40	21	2.8	3701	9	AF080617 Homo sapi
41	21	2.8	4935	12	AF050499 Cloning v
42	21	2.8	4935	12	AF063584 Cloning v
43	21	2.8	6649	9	AF080618 Homo sapi
44	21	2.8	6847	9	AF080615 Homo sapi
45	21	2.8	10558	9	AB016195 Homo sapi

ALIGNMENTS

RESULT 1
AX323453 LOCUS AX323453 759 bp DNA linear PAT 07-JAN-2002
DEFINITION Sequence 1 from Patent WO0192308.
ACCESSION AX323453
VERSION AX323453.1 GI:18094216
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Paszty,C.J. and Gao,Y.
TITLE Cysteine-knot polypeptides: cloaked-2 molecules and uses thereof
JOURNAL Patent: WO 0192308-A 1 06-DEC-2001;

QY 421 GGGCCGCACTTCGCTGATCCCGACCGCTACCGCGCGAGCGCGT-2AGTGTGTGT 480
 DB 422 GGGCCGCACTTCGCTGATCCCGACCGCTACCGCGCGAGCGCGT-2AGTGTGT 481
 QY 481 CCCGCTGCTAGCGCGCGCGCGCGCGCAAGGTGCGCTGTGTGTGTGTGTGTGTGT 540
 DB 482 CCCGCTGCTAGCGCGCGCGCGCGCGCAAGGTGCGCTGTGTGTGTGTGTGTGTGT 541
 QY 541 CGGCTACCGCGCTTCCACACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
 DB 542 CGGCTACCGCGCTTCCACACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 601
 QY 601 CCGCAGAGAGCGCGAGAGCGCGCGCGCGCGCGCGAGCGCAAGCGCAAGCGCGAG 660
 DB 602 CCGCAGAGAGCGCGAGAGCGCGCGCGCGCGCGCGAGCGCAAGCGCAAGCGCGAG 661
 QY 661 CTGAGAGAGCGCTTCTAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 720
 DB 662 CTGAGAGAGCGCTTCTAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 721
 QY 721 GAACCGCGCGCGCGCAATTTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 759
 DB 722 GAACCGCGCGCGCGCAATTTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 760

RESULT 3
 AF326739 2323 bp mRNA linear PRI 28-FEB-2001
 LOCUS Homo sapiens sclerostin mRNA, complete cds.
 DEFINITION AF326739
 ACCESSION AF326739 GI:13161019
 VERSION
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
 AUTHORS 1 (bases 1 to 2323)
 Brunkow, M.E., Gardner, J.C., Van Ness, J., Paepker, B.W.,
 Kovacevich, B.R., Prohl, S., Skonier, J.E., Zhao, Y., Sabo, P.J.,
 Fu, Y.H., Altsch, R.S., Gillett, L., Colbert, T., Tacconi, P., Galas, D.,
 Hamersma, H., Beighton, P. and Mulligan, J.T.
 Bone dysplasia sclerosteosis results from loss of the SOST gene
 product, a novel cysteine knot-containing protein
 Am. J. Hum. Genet. 68 (3), 577-589 (2001)

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS

2 (bases 1 to 2323)
 Brunkow, M.E., Gardner, J.C., Van Ness, J., Paepker, B.W.,
 Kovacevich, B.R., Prohl, S., Skonier, J.E., Zhao, Y., Sabo, P.J.,
 Fu, Y.H., Altsch, R.S., Gillett, L., Colbert, T., Tacconi, P., Galas, D.,
 Hamersma, H., Beighton, P. and Mulligan, J.T.
 Direct Submission
 Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631
 220th St. SE, Bothell, WA 98021, USA

FEATURES
 source
 1. 2323
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="17"
 /map="17q21"
 1. 47
 48..689
 /codon_start=1
 /product="sclerostin"
 /protein_id="AAK13454.1"
 /translation="MQLPLALCLVCLVHTAFVVEGQGMQAFKNDATETIIPELGEYP
 EPPLENNKTKMARANGRPHPHPETDVSYSRSHFRVYTTDPCRSAPKYTE
 LVSCGCGPARLPNALIGRKMWRSGPFCIPRYRZKORVOLLCPGGEAPRAKVR
 LVASCKRLLTFHNSSELDKDFTEAARQKRRPRPARSKANQAELENNY"
 690..2323

3'UTR
 BASE COUNT
 590 a 635 c 614 g 484 t

ORIGIN

Query Match 100.0%; Score 759; DB 9; Length 2323;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 759; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TACTGAAAGTGGCGTGGCTTCTCTGTGCTGATCATGAGCTTCCACTGCGCTGTGT 60
 DB 12 TACTGAAAGTGGCGTGGCTTCTCTGTGCTGATCATGAGCTTCCACTGCGCTGTGT 71
 QY 61 CTGCTGTGCTGCTGT 120
 DB 72 CTGCTGTGCTGCTGT 131
 QY 121 TTCAAGAAATGATGACCAAGGAAATCATCCCGAGCTGGAGAGTACCCGAGCTTCAACG 180
 DB 132 TTCAAGAAATGATGACCAAGGAAATCATCCCGAGCTGGAGAGTACCCGAGCTTCAACG 191
 QY 181 GAGCTGAGAAACAAACAGACCAATGAAACCGGGGAGAAAGAGGGCGCTTCCACACAC 240
 DB 192 GAGCTGAGAAACAAACAGACCAATGAAACCGGGGAGAAAGAGGGCGCTTCCACACAC 251
 QY 241 CCTTTGAGAACCAAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
 DB 252 CCTTTGAGAACCAAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 311
 QY 301 GTGACGATGGGCGGTGCGGAGCGCGCAAGCGGTGCAACGAGCTGTGTGTGTGTGTGT 360
 DB 312 GTGACGATGGGCGGTGCGGAGCGCGCAAGCGGTGCAACGAGCTGTGTGTGTGTGTGT 371
 QY 361 TCGGCGCGCGCGCGCTGT 420
 DB 372 TCGGCGCGCGCGCGCTGT 431
 QY 421 GGGCCGCACTTCGCTGATCCCGACCGCTACCGCGCGAGCGCGTCACTTCAACCGCTAC 480
 DB 432 GGGCCGCACTTCGCTGATCCCGACCGCTACCGCGCGAGCGCGTCACTTCAACCGCTAC 491
 QY 481 CCGGCTGTGAGAGCGCGCGCGCGCGCGCAAGGTGCGCTGTGTGTGTGTGTGTGTGTGT 540
 DB 492 CCGGCTGTGAGAGCGCGCGCGCGCGCGCAAGGTGCGCTGTGTGTGTGTGTGTGTGT 551
 QY 541 CGCTTACCGCTTCCACACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
 DB 552 CGCTTACCGCTTCCACACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 611
 QY 601 CCGCAGAGAGCGCGAGAGCGCGCGCGCGCGCGCGAGCGCGCAAGCGCAAGCGCGAG 660
 DB 612 CCGCAGAGAGCGCGAGAGCGCGCGCGCGCGCGCGAGCGCGCAAGCGCAAGCGCGAG 671
 QY 661 CTGAGAGAGCGCTTCTAGAGAGCGCGCGCGCGCGCGCGAGCGCGCAAGCGCGAG 720
 DB 672 CTGAGAGAGCGCTTCTAGAGAGCGCGCGCGCGCGCGCGAGCGCGCAAGCGCGAG 731
 QY 721 GAACCGCGCGCGCGCAATTTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 759
 DB 732 GAACCGCGCGCGCGCAATTTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 760

RESULT 4
 AX056687 2329 bp DNA linear PAT 17-JAN-2001
 LOCUS Sequence 19 from Patent WO0075317.
 DEFINITION AX056687
 ACCESSION AX056687
 VERSION
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
 AUTHORS 1 (bases 1 to 2329)
 Borstein, D.A., Goddard, A., Gurney, A.L., Smith, V., Watanabe, C.K. and
 Wood, W.I.

Db 721 CTGCGCGGTGTT 725

RESULT 6
LOCUS AX342537/c
DEFINITION Sequence 3 from Patent WO0198491.
ACCESSION AX342537
VERSION AX342537.1 GI:18151965
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Bailemans, W., Ebeling, M., Foenzler, D., Patel, N., van Hul, W. and Vickers, B.H.
TITLE Osteolevin gene polymorphisms
JOURNAL Patent: WO 0198491-A 3 27-DEC-2001, F. HOFMANN-LA ROCHE AG (CH) ; UNIVERSITAIRE INSTELLING ANTWERPEN (BE)

FEATURES
source Location/Qualifiers
1..7099
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 1530 a 1944 c 1928 g 1697 t

ORIGIN

Query Match 66.5%; Score 505; DB 6; Length 7099;
Best Local Similarity 100.0%; Pred. No. 6e-242;
Matches 505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 255 AGAGGTGTCCAGTACAGCTGCGCGAGCTGCACTTCAACCGCTACCGATGGGCC 314
Db 2343 AGAGGTGTCCAGTACAGCTGCGCGAGCTGCACTTCAACCGCTACCGATGGGCC 2284

QY 315 GTGCGCGAGCG 374
Db 2283 GTGCGCGAGCG 2224

QY 375 CTTGCTCCCAACCG 434
Db 2223 CTTGCTCCCAACCG 2164

QY 435 CTGCAATCCCGACCG 494
Db 2163 CTGCAATCCCGACCG 2104

QY 495 GCG 554
Db 2103 GCG 2044

QY 555 CCAACAACAGTGGAGTCAAGGACTTGGGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCG 614
Db 2043 CCAACAACAGTGGAGTCAAGGACTTGGGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCG 1984

QY 615 GAAAGCG 674
Db 1983 GAAAGCG 1924

QY 675 CTGAGAGCG 734
Db 1923 CTGAGAGCG 1864

QY 735 CATTTCTGTCTCTGCGCGGTGTT 759
Db 1863 CATTTCTGTCTCTGCGCGGTGTT 1839

RESULT 7
LOCUS AF326736
DEFINITION Homo sapiens sclerostin gene, complete cds.

ACCESSION AF326736
VERSION AF326736.1 GI:13161010
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Brunkow, M.E., Gardner, J.C., Van Ness, J., Paepel, B.W., Kovacevich, B.R., Proll, S., Skonier, J.E., Zhao, L., Sabo, P.J., Fu, Y.H., Alisch, R.S., Gillett, L., Colbert, T., Tacconi, P., Galas, D., Hamersma, H., Beighton, P. and Mulligan, J.T.
TITLE Bone dysplasia sclerostosis results from loss of the SOST gene product, a novel cysteine knot-containing protein
JOURNAL Am. J. Hum. Genet. 68 (3), 577-589 (2001)
MEDLINE 21090529
PUBMED 11179006

REFERENCE
PUBLISHED 2 (bases 1 to 21501)
AUTHORS Brunkow, M.E., Gardner, J.C., Van Ness, J., Paepel, B.W., Kovacevich, B.R., Proll, S., Skonier, J.E., Zhao, L., Sabo, P.J., Fu, Y.H., Alisch, R.S., Gillett, L., Colbert, T., Tacconi, P., Galas, D., Hamersma, H., Beighton, P. and Mulligan, J.T.
TITLE Direct Submission
JOURNAL Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631 220th St. SE, Bothell, WA 98021, USA

FEATURES
source Location/Qualifiers
1..21501
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q21"
join(111961..12180,14939..15360)
/product="sclerostin"
join(111961..12180,14939..15360)
/codon_start=1
/product="sclerostin"
/protein_id="AAK13451.1"
/db_xref="GI:13161011"
translation="MQPLALCLVCLVHTAFRVVGGGQAFNDAETIIPEIGXP
BPPELENNKTMRAENGSRPBNHPRFEDKDYSEVSCPLHETRYVTGPGSARPYTE
LVSCQCGPALLPNAIGRGKWRPSSGDFPCIDPRDAQGVQLCPGEGAPRKAKVR
LVASCKRLTRFNNQSELDKGTGAAPRQGRKPRPARSAKANQAELENAY"

BASE COUNT 5546 a 5259 c 5419 g 5277 t

ORIGIN

Query Match 66.5%; Score 505; DB 9; Length 21501;
Best Local Similarity 100.0%; Pred. No. 5.9e-242;
Matches 505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 255 AGAGGTGTCCAGTACAGCTGCGCGAGCTGCACTTCAACCGCTACCGATGGGCC 314
Db 14937 AGAGGTGTCCAGTACAGCTGCGCGAGCTGCACTTCAACCGCTACCGATGGGCC 14996

QY 315 GTGCGCGAGCG 374
Db 14997 GTGCGCGAGCG 15056

QY 375 CTTGCTCCCAACCG 434
Db 15057 CTTGCTCCCAACCG 15116

QY 435 CTGCAATCCCGACCG 494
Db 15117 CTGCAATCCCGACCG 15176

QY 495 GCG 554
Db 15177 GCG 15236

QY 555 CCAACAACAGTGGAGTCAAGGACTTGGGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCG 614
Db 15237 CCAACAACAGTGGAGTCAAGGACTTGGGAGCGGAGCGCGCGCGCGCGCGCGCGCGCGCG 15296

repeat_region complement (2492. .2787)
 /rpt_family="AluSg"
 repeat_region complement (2796. .2902)
 /rpt_family="MIR"
 repeat_region complement (2903. .3165)
 /rpt_family="AluSg"
 repeat_region complement (3166. .3193)
 /rpt_family="MIR"
 repeat_region 3572. .3644
 /rpt_family="MIR"
 repeat_region 4524. .4583
 /rpt_family="MIR3"
 repeat_region complement (4755. .4962)
 /rpt_family="MIR"
 repeat_region 6397. .6738
 /rpt_family="L2"
 repeat_region complement (6766. .6875)
 /rpt_family="MLT1H"
 repeat_region 7202. .7270
 /rpt_family="L2"
 repeat_region 7739. .7825
 /rpt_family="L1MC4"
 repeat_region complement (7826. .8122)
 /rpt_family="AluSc"
 repeat_region 8123. .8268
 /rpt_family="L1MC4"
 repeat_region 8269. .8568
 /rpt_family="AluJo"
 repeat_region 8569. .8631
 /rpt_family="L1MC4"
 repeat_region complement (9151. .9305)
 /rpt_family="MIR"
 repeat_region 9365. .9390
 /rpt_family="AT rich"
 repeat_region complement (9405. .9708)
 /rpt_family="AluY"
 repeat_region complement (9709. .9969)
 /rpt_family="AluJo"
 repeat_region 10197. .10458
 /rpt_family="MLT1K"
 repeat_region 10758. .10895
 /rpt_family="MIR"
 repeat_region 12676. .12773
 /rpt_family="MIR"
 repeat_region complement (12784. .12997)
 /rpt_family="MIR"
 repeat_region complement (13220. .13513)
 /rpt_family="AluJo"
 repeat_region 13627. .13787
 /rpt_family="L2"
 repeat_region complement (13821. .13953)
 /rpt_family="MIR"
 repeat_region 14746. .14899
 /rpt_family="MIR"
 repeat_region complement (14947. .15144)
 /rpt_family="MIR"
 repeat_region 15520. .15668
 /rpt_family="AluJo"
 repeat_region 15665. .15960
 /rpt_family="AluSx"
 repeat_region 15961. .16128
 /rpt_family="AluJo"
 repeat_region 16208. .16400
 /rpt_family="MIR"
 repeat_region 16454. .16586
 /rpt_family="L2"
 repeat_region 16707. .16783
 /rpt_family="MIR"
 repeat_region 16852. .16909

Query Match 66.5% Score 505 DB 9 Length 80117
 Best Local Similarity 100.0% Fred No. 5.8e-242
 Matches 505 Conservative 0 Mismatches 0 Indels 0 Gaps 0

QY 255 AGACGTCGACGAGTCGCGGAGCTGACCTTCAACCCGCTACGTCGATGAGGCC 314
 |||||
 Db 63216 AGACGTCGACGAGTCGCGGAGCTGACCTTCAACCCGCTACGTCGATGAGGCC 63157
 |||||
 QY 315 GTGCGCAGCGCAGCGCAGCGCAGCGCAGCGCAGCGCAGCGCAGCGCAGCGC 374
 |||||
 Db 63156 GTGCGCAGCGCAGCGCAGCGCAGCGCAGCGCAGCGCAGCGCAGCGCAGCGC 63097
 |||||
 QY 375 CTGCTGCTCCCAACGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 434
 |||||
 Db 63096 CTGCTGCTCCCAACGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 63037
 |||||
 QY 435 CTGCTGCTCCCAACGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 494
 |||||
 Db 63036 CTGCTGCTCCCAACGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 62977
 |||||
 QY 495 GCG 554
 |||||
 Db 62976 GCG 62917
 |||||
 QY 555 CCACACACGTCGAGCTTCAAGAGCTTGGGAGCGAGCGCGCGCGCGCGCGCGCGCG 614
 |||||
 Db 62916 CCACACACGTCGAGCTTCAAGAGCTTGGGAGCGAGCGCGCGCGCGCGCGCGCG 62857
 |||||
 QY 615 GAAGCG 674
 |||||
 Db 62856 GAAGCG 62797
 |||||
 QY 675 CTAGAGCCCG 734
 |||||
 Db 62796 CTAGAGCCCG 62737
 |||||
 QY 735 CATTTCTGCT 759
 |||||
 Db 62736 CATTTCTGCT 62712
 |||||

RESULT 9

AC003098/c 94752 bp DNA linear PRI 29-JAN-1998
 LOCUS Homo sapiens chromosome 17, clone HRP0905N1, complete sequence.
 AC003098
 VERSION AC003098.1 GI:2822155
 KEYWORDS HNG.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 Birren, B., Pasman, K., McKernan, K., Nusbaum, C., Richardson, P. and Lander, E. (bases 1 to 94752)
 2 (bases 1 to 94752)
 TITLE Homo sapiens chromosome 17, clone HRP0905N1
 JOURNAL Unpublished
 REFERENCE AUTHORS

TITLE Direct Submission
 JOURNAL Submitted (10-NOV-1997) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 REFERENCE 3 (bases 1 to 94752)
 AUTHORS Birren, B., Pasman, K., McKernan, K., Nusbaum, C., Richardson, P., Lander, E., Allen, A., Baker, J., Baldwin, J., Barina, N., Beckert, R., Boutwell, C., Byrne, S., Cantu, C., Castle, A., Cooke, P., Daly, M., J., Depayre, B., Devon, K., Dewar, K., Dubette, B., Etemadi, S.,


```

repeat_region      /rpt_family="(GA)n"
                    complement(35561..35673)
repeat_region      /rpt_family="MLT2D"
                    complement(35678..35935)
repeat_region      /rpt_family="MER66B"
                    complement(35936..36233)
repeat_region      /rpt_family="Alus9"
                    complement(36236..36449)
repeat_region      /rpt_family="MER66B"
                    complement(36450..36867)
repeat_region      /rpt_family="MLT2D"
                    complement(36876..37460)
repeat_region      /rpt_family="L1MD2"
                    complement(37477..37777)
repeat_region      /rpt_family="AlusX"
                    complement(37804..37948)
repeat_region      /rpt_family="L1"
                    complement(38130..38175)
repeat_region      /rpt_family="L1MD1"
                    complement(38253..38253)
repeat_region      /rpt_family="L1MD2"

```

```

Query Match      66.5%; Score 505; DB 9; Length 94752;
Best Local Similarity 100.0%; Pred. No. 5 Be-242;
Matches 505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 255 AGAGCTGTCGCAAGTACAGCTCCCGGAGCTGCACTTCAACCGCTACCGATGAGGCC 314
Db 6293 AGAGCTGTCGCAAGTACAGCTCCCGGAGCTGCACTTCAACCGCTACCGATGAGGCC 6234
Qy 315 GTGCGCAGCGCCAGAGCGGTACCGAGCTGCTGCTCCGCGCAAGTGCAGCGCGCGCG 374
Db 6233 GTGCGCAGCGCCAGAGCGGTACCGAGCTGCTGCTCCGCGCAAGTGCAGCGCGCGCG 6174
Qy 375 CTTGCTGTCGCAAGTACAGCTCCCGGAGCTGCACTTCAACCGCTACCGATGAGGCC 434
Db 6173 CTTGCTGTCGCAAGTACAGCTCCCGGAGCTGCACTTCAACCGCTACCGATGAGGCC 6114
Qy 435 CTGATATCCCGGAGCGGTACCGAGCTGCTGCTCCGCGCAAGTGCAGCGCGCGCGCG 494
Db 6113 CTGATATCCCGGAGCGGTACCGAGCTGCTGCTCCGCGCAAGTGCAGCGCGCGCGCG 6054
Qy 495 GCCGCGCGCGCGAGGTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 554
Db 6053 GCCGCGCGCGCGAGGTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5994
Qy 555 CCAGAACCAAGTCGAGCTCAAGGACCTTGGGACCGAGCGCTGCGCCGCAAGAGGCCG 614
Db 5993 CCAGAACCAAGTCGAGCTCAAGGACCTTGGGACCGAGCGCTGCGCCGCAAGAGGCCG 5934
Qy 615 GAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 674
Db 5933 GAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5874
Qy 675 CTAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 734
Db 5873 CTAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 5814
Qy 735 CATTTCTGCTCTTCTGCGCGTGTGTT 759
Db 5813 CATTTCTGCTCTTCTGCGCGTGTGTT 5789

```

```

RESULT 10
AC073954 177744 bp DNA linear HTG 17-NOV-2000
LOCUS AC073954 Homo sapiens chromosome 17 clone RP11-147C10, WORKING DRAFT
DEFINITION AC073954
ACCESSION AC073954
VERSION AC073954.3 GI:10998993
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens.
ORGANISM Homo sapiens

```

REFERENCE AUTHORS TITLE JOURNAL REFERENCE AUTHORS TITLE JOURNAL

COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 177744) Waterston, R.H. The sequence of Homo sapiens clone 2 (bases 1 to 177744) Waterston, R.H. Direct Submission Submitted (07-JUL-2000) Genome Sequencing Center, Washington University School of Medicine, 444 Forest Park Parkway, St. Louis, MO 63108, USA On Oct 25, 2000 this sequence version replaced gi:1986649.

Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>
Project Information
Center project name: H.NH0147C10

Summary Statistics
Sequencing vector: M13; 100%
Sequencing method: plasmid; 0%
Chemistry: Dye-terminator Big Dye; 0% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 155679 bases at least Q40
Consensus quality: 163329 bases at least Q30
Insert size: 185000; agarose-fp
Insert size: 183559; sum-of-contigs
Quality coverage: 3.29 in Q20 bases; agarose-fp
Quality coverage: 3.43 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 34 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence, as soon as it is available and the accession number will be preserved.

1	1247: contig of 1247 bp in length
1248	1347: gap of unknown length
1348	2762: contig of 1415 bp in length
2763	2862: gap of unknown length
2863	4773: contig of 1911 bp in length
4774	4873: gap of unknown length
4874	6326: contig of 1453 bp in length
6327	6426: gap of unknown length
6427	8873: contig of 2447 bp in length
8874	8973: gap of unknown length
8974	10783: contig of 1810 bp in length
10784	10883: gap of unknown length
10884	12763: contig of 1880 bp in length
12764	12863: gap of unknown length
12864	15381: contig of 2518 bp in length
15382	15481: gap of unknown length
15482	17365: contig of 1884 bp in length
17366	17465: gap of unknown length
17466	19992: contig of 2527 bp in length
19993	20092: gap of unknown length
20093	23512: contig of 3420 bp in length
23513	23612: gap of unknown length
23613	25699: contig of 2087 bp in length
25700	25799: gap of unknown length
25800	28333: contig of 2534 bp in length
28334	28433: gap of unknown length
28434	32027: contig of 3594 bp in length
32028	32127: gap of unknown length
32128	36115: contig of 3988 bp in length
36116	36215: gap of unknown length
36216	39882: contig of 3667 bp in length

REFERENCE
AUTHORS
1 (bases 1 to 642)
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Cercopithecus.
Brunkow, M.E., Gardner, J.C., Van Ness, J., Paepers, B.W., Kovacevich, B.R., Prohl, S., Skonier, J.E., Zhao, L., Sabo, P.J., Fu, Y.H., Alisch, R.S., Gillett, L., Colbert, T., Tacconi, P., Galas, D., Hamersma, H., Beighton, P. and Mulligan, J.T.
Bone dysplasia scleroosteosis results from loss of the SOST gene product, a novel cysteine knot-containing protein
Am. J. Hum. Genet. 68 (3), 577-589 (2001)

JOURNAL
MEDLINE
PUBMED
21090529
11179006

REFERENCE
AUTHORS
2 (bases 1 to 642)
Brunkow, M.E., Gardner, J.C., Van Ness, J., Paepers, B.W., Kovacevich, B.R., Prohl, S., Skonier, J.E., Zhao, L., Sabo, P.J., Fu, Y.H., Alisch, R.S., Gillett, L., Colbert, T., Tacconi, P., Galas, D., Hamersma, H., Beighton, P. and Mulligan, J.T.
Direct Submission
Submitted (07-DEC-2000) Genomics, Celtech Chiroscience Inc., 1631 220th St. SE, Bothell, WA 98021, USA

TITLE
JOURNAL
220th St. SE, Bothell, WA 98021, USA

FEATURES
source
location/Qualifiers
1..642
/organism="Cercopithecus aethiops"
/db_xref="taxon:9534"
1..642
/product="sclerostin"
/protein_id="AAK13457.1"
/db_xref="GI:13161029"
/translation="MDPLALCLVCLVHAAPVVEGCGMAFKNDATITPELGEYP
EPPELENNKTRNAENGRRPHHPPTDVSYSRELFHFRYVTDGCRSAKVT
LVSGCGGPARLPNALIGKMRSPGPRCPIDYRKRORVLLCPGSAAPRAKVR
LVASCKKRLTRHNOSELDPFPAARPKPKPRARAKAQAQALENNY"

CDS
BASE COUNT
113 a 228 c 210 g 91 t

Query Match
Best Local Similarity 14.5%; Score 110; DB 9; Length 642;
Matches 210; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db
121 TTCAAGATGATGCGCAAGAAATATCTCCGAGCTGGAGAGTACCCGAGCTCCACCG 180
85 TTCAAGATGATGCGCAAGAAATATCTCCGAGCTGGAGAGTACCCGAGCTCCACCG 144

Qy
181 GAGCTGGAACAACAACAAGACCATGATGAGCTGGCGGAGCTGACCTTCCACCG 240
145 GAGCTGGAACAACAACAAGACCATGATGAGCTGGCGGAGCTGACCTTCCACCG 204

Db
241 CCTTTGAGACCAAGAAGCGTGTCCGAGTACAGCTGGCGGAGCTGACCTTCCACCG 300
205 CCTTTGAGACCAAGAAGCGTGTCCGAGTACAGCTGGCGGAGCTGACCTTCCACCG 264

Qy
301 GTGACCGATGGCG 332
265 GTGACCGATGGCG 296

RESULT 12
AF326738 532 bp mRNA linear MAM 28-FEB-2001
LOCUS
DEFINITION
Bos taurus sclerostin mRNA, partial cds.
AF326738.1 GI:13161016
KEYWORDS
SOURCE
ORGANISM
Bos taurus.
Bos taurus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.
REFERENCE
AUTHORS
1 (bases 1 to 532)
Brunkow, M.E., Gardner, J.C., Van Ness, J., Paepers, B.W., Kovacevich, B.R., Prohl, S., Skonier, J.E., Zhao, L., Sabo, P.J., Fu, Y.H., Alisch, R.S., Gillett, L., Colbert, T., Tacconi, P., Galas, D.,

TITLE
JOURNAL
MEDLINE
PUBMED
21090529
11179006

REFERENCE
AUTHORS
2 (bases 1 to 532)
Brunkow, M.E., Gardner, J.C., Van Ness, J., Paepers, B.W., Kovacevich, B.R., Prohl, S., Skonier, J.E., Zhao, L., Sabo, P.J., Fu, Y.H., Alisch, R.S., Gillett, L., Colbert, T., Tacconi, P., Galas, D., Hamersma, H., Beighton, P. and Mulligan, J.T.
Direct Submission
Submitted (07-DEC-2000) Genomics, Celtech Chiroscience Inc., 1631 220th St. SE, Bothell, WA 98021, USA

TITLE
JOURNAL
220th St. SE, Bothell, WA 98021, USA

FEATURES
source
location/Qualifiers
1..532
/organism="Bos taurus"
/db_xref="taxon:9913"
1..532
/product="sclerostin"
/protein_id="AAK13453.1"
/db_xref="GI:13161017"
/translation="NDATETIPELGEYEPPELENNKTRNAENGRRPHHPPTD
SEVSCRELFHFRYVTDGCRSAKVTPELVSGCGGPARLPNALIGKMRSPGPR
CIDPRVAORVOLLCPGSAAPRAKVRKRLVASCCKRLTRHNOSELKDFPAPARQT
GRKLRPRARAKTSRA"

CDS
BASE COUNT
95 a 199 c 176 g 62 t

Query Match
Best Local Similarity 10.9%; Score 83; DB 4; Length 532;
Matches 133; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db
280 GAGCTGCACTTACCCGCTACGATGAGCGCGCTGGCGGAGCGCAAGCGGTACCG 339
153 GAGCTGCACTTACCCGCTACGATGAGCGCGCTGGCGGAGCGCGCAAGCGGTACCG 212

Qy
340 GAGCTGATGCTCTCGGCGCAAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 399
213 GAGCTGATGCTCTCGGCGCAAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 272

Db
400 GCGAAGTGTGGCG 413
273 GCGAAGTGTGGCG 286

RESULT 13
AF397423 93790 bp DNA linear PRI 02-APR-2002
LOCUS
DEFINITION
Homo sapiens SOST gene, 3' UTR, MBOX1 gene, 5' UTR, and SOST/MBOX1 intergenic region.
AF397423
AF397423.1 GI:19880618
KEYWORDS
SOURCE
ORGANISM
Homo sapiens.
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
AUTHORS
1 (bases 1 to 93790)
Staeinling-Hampton, K., Prohl, S., Paepers, B., Zhao, L., Charney, P., Brown, A., Gardner, J.C., Galas, D., Schatzman, R.C., Beighton, P., Papapoulos, S., Hamersma, H. and Brunkow, M.E.
A 52 kb deletion in the SOST - MBOX1 intergenic region on 17q12-q21 is associated with van Buchem disease in the Dutch population
Unpublished
2 (bases 1 to 93790)
Brunkow, M., Staeinling-Hampton, K., Prohl, S., Paepers, B., Zhao, L., Charney, P., Brown, A., Gardner, J.C., Galas, D., Schatzman, R., Beighton, P., Papapoulos, S. and Hamersma, H.
Direct Submission
Submitted (06-JUN-2001) Genomics, Celtech R&D Inc., 1631 220th St SE, Bothell, WA 98021, USA

FEATURES

Location/Qualifiers

1.33790

source

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="17"

/map="17q21-q22"

<1.1612

/gene="SOST"

<1.1612

/gene="SOST"

mRNA

3' UTR

1.1612

/gene="SOST"

12298.12676

/standard_name="D17S1789"

36239.87958

/note="Van Buchem causative deletion"

variation

STS

79495.79823

/standard_name="D17S1789"

93762.93790

/gene="MEOX1"

93762.93790

/gene="MEOX1"

93762.93790

/gene="MEOX1"

5' UTR

93762.93790

/gene="MEOX1"

BASE COUNT 24072 a 23664 c 22642 g 23409 t 3 others

ORIGIN

Query Match 10.7%; Score 81; DB 9; Length 93790;

Best Local Similarity 100.0%; Pred. No. 2.2e-29; Indels 0; Gaps 0;

Matches 81; Conservative 0; Mismatches 0;

OY 679 AGCCGCGCCGCGCCCTCCCAACGCGCGCGCGCCGCGCTGAACCCGCGCCCAATT 738

DB 1 AGCCGCGCCGCGCCCTCCCAACGCGCGCGCGCCGCGCTGAACCCGCGCCCAATT 60

OY 739 TCTGTCTCTGCGCGTGT 759

DB 61 TCTGTCTCTGCGCGTGT 81

RESULT 14

AX323455 636 bp DNA linear PAT 07-JAN-2002

LOCUS AX323455 3 from Patent WO0192308.

DEFINITION Sequence

ACCESSION AX323455

VERSION AX323455.1 GI:18094217

KEYWORDS

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1

Paszty,C.J. and Gao,Y.

Cystine-knot polypeptides: cloaked-2 molecules and uses thereof

Patent: WO 0192308-A 3 06-DEC-2001;

JOURNAL Amgen, Inc. (US)

FEATURES

Location/Qualifiers

1.636

/organism="Mus musculus"

/db_xref="taxon:10090"

BASE COUNT 114 a 224 c 207 g 91 t

ORIGIN

Query Match 9.1%; Score 69; DB 6; Length 636;

Best Local Similarity 100.0%; Pred. No. 2.3e-23; Indels 0; Gaps 0;

Matches 69; Conservative 0; Mismatches 0;

OY 517 CTGTGGCTGTGAGTGAAGGCTCAACCGCTTCCACAACAGTGGAGCTCAAG 576

DB 475 CTGTGGCTGTGAGTGAAGGCTTCCACAACAGTGGAGCTCAAG 534

OY 577 GACTTCGGG 585

DB 535 GACTTCGGG 543

RESULT 15

AF326740 638 bp mRNA linear ROD 28-FEB-2001

LOCUS AF326740

DEFINITION Mus musculus sclerostin mRNA, complete cds.

ACCESSION AF326740

VERSION AF326740.1 GI:13161022

KEYWORDS

SOURCE

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1

Brunkow,M.E., Gardner,J.C., Van Ness,J., Paeppe,B.W., Kovacevich,B.R., Prohl,S., Skonier,J.E., Zhao,L., Sabo,P.J., Fu,Y.H., Alisch,R.S., Gillett,L., Colbert,T., Tacconi,P., Galas,D., Hamersma,H., Beighton,P. and Mulligan,J.T.

Bone dysplasia sclerosteosis results from loss of the SOST gene product, a novel cystine knot-containing protein

Am. J. Hum. Genet. 68 (3), 577-589 (2001)

JOURNAL

MEDLINE

PUBMED

21090529

11179006

2 (bases 1 to 638)

Brunkow,M.E., Gardner,J.C., Van Ness,J., Paeppe,B.W., Kovacevich,B.R., Prohl,S., Skonier,J.E., Zhao,L., Sabo,P.J., Fu,Y.H., Alisch,R.S., Gillett,L., Colbert,T., Tacconi,P., Galas,D., Hamersma,H., Beighton,P. and Mulligan,J.T.

Direct Submission

Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631 220th St. SE, Bothell, WA 98021, USA

FEATURES

Location/Qualifiers

1.638

/organism="Mus musculus"

/strain="129Sv/J"

/db_xref="taxon:10090"

/chromosome="11"

1.636

/codon_start=1

/product="sclerostin"

/protein_id="AAK13455.1"

/db_xref="GI:13161023"

/translation="MQPSLAPPCILVHAAPGAVGQGMQAFRNDATVPIGLGEYP EPPENNQTMRANIGRPPHPYDIAKGVSEVSCRLHYTRRLTGGCSARVTELV CSQCGQRLPLPAIGIKVXWRNGDPFCTIPRYRQVQLCRGAA PRARKYRLV ASCKRKLRTFHNSSEIKDPGPEPTARPKGRKRPGRKAKANQAELENAY"

BASE COUNT 114 a 224 c 209 g 91 t

ORIGIN

Query Match 9.1%; Score 69; DB 10; Length 638;

Best Local Similarity 100.0%; Pred. No. 2.3e-23; Indels 0; Gaps 0;

Matches 69; Conservative 0; Mismatches 0;

OY 517 CTGTGGCTGTGAGTGAAGGCTCAACCGCTTCCACAACAGTGGAGCTCAAG 576

DB 475 CTGTGGCTGTGAGTGAAGGCTTCCACAACAGTGGAGCTCAAG 534

OY 577 GACTTCGGG 585

DB 535 GACTTCGGG 543

Search completed: March 29, 2003, 09:56:49

Job time: 2587.25 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 28, 2003, 23:23:29 ; Search time 237.766 Seconds
(without alignments)
7188.873 Million cell updates/sec

Title: US-09-867-274-1

Perfect score: 759
Sequence: 1 tctcgaaggtggtgtcc.....ctgtcctctgtgtgttt 759

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	759	100.0	759	24	ABA94293
2	759	100.0	2301	21	AAA29055
3	757	99.7	2329	22	AAA91023
4	757	99.7	2329	24	ABK69992
5	735	96.8	2371	24	ABD27576
6	708	93.3	2301	21	AAA29061
7	708	93.3	2301	21	AAA29062
8	657	86.6	2301	21	AAA29056
9	642	84.6	642	21	AAA94051

10	505	66.5	5680	21	AAA94049	Human DAN/Cerberus
11	505	66.5	7099	24	AAD27577	Human osteolevin g
12	505	66.5	9301	21	AAA29064	Human TGF-beta bin
13	422	55.6	422	22	AAK45509	Human bone marrow
14	422	55.6	422	24	AB519777	Human genome-deriv
15	422	55.6	1104	21	AAA94050	Human DAN/Cerberus
16	355	46.8	392	22	AAK32487	Human bone marrow
17	355	46.8	392	22	AB507272	Human genome-deriv
18	182	24.0	196	21	AAK05741	Human secreted pro
19	182	24.0	198	20	AAK40842	Human secreted pro
20	110	14.5	642	21	AAA29057	Veret TGF-beta bi
21	83	10.9	532	24	AAA29060	Bovine TGF-beta bi
22	69	9.1	636	24	ABA94294	Mouse cloaked-2 po
23	69	9.1	638	21	AAA29058	Murine TGF-beta bi
24	69	9.1	35828	21	AAA29063	Murine TGF-beta bi
25	68	9.0	674	21	AAA29059	Human secreted pro
26	40	5.3	40	22	AAA91036	Human secreted pro
27	28	3.7	39	21	AAA94042	Rat TGF-beta bindi
28	28	3.7	54	21	AAA94041	Probe for Human se
29	27	3.6	27	22	AAA91035	Human hDCR6 exon 1
30	27	3.6	27	24	ABA94309	Human hDCR6 exon 1
31	27	3.6	27	22	AAA94309	PCR primer for Hum
32	26	3.4	38	21	AAA94048	Human cloaked-2 co
33	26	3.4	26	24	ABA94296	Human hDCR6 exon 4
34	25	3.3	30	21	AAA94047	Human hDCR6 exon 4
35	24	3.2	39	21	AAA29072	Human hDCR6 exon 4
36	24	3.2	24	24	ABA94295	Sense primer to am
37	24	3.2	555	24	ABQ35254	Human cloaked-2 cd
38	23	3.0	555	24	ABQ35255	Oligonucleotide fo
39	23	3.0	23	24	ABA94310	Oligonucleotide fo
40	21	2.8	21	24	AAD27580	Human cloaked-2 co
41	21	2.8	21	21	AAA29065	Human osteolevin g
42	21	2.8	21	22	AAA91034	Sense PCR primer f
43	21	2.8	41	24	ABA94308	PCR primer for Hum
44	21	2.8	57	21	AAA29073	Mouse cloaked-2 cd
45	21	2.8	2266	20	AAZ41049	Antisense primer t
			2266	20	AAZ06608	Human ERK-1 nucleo

ALIGNMENTS

RESULT 1	
ID	ABA94293 standard; cDNA: 759 BP.
XX	XX
AC	ABA94293;
XX	XX
DT	26-MAR-2002 (first entry)
XX	XX
DE	Human cloaked-2 polypeptide encoding cDNA.
XX	XX
XX	Cloaked-2: cysteine knot motif; nephrotropic; cardiant; immunomodulator;
KW	hepatotropic; antiinflammatory; antithyroid; cytostatic; neuroprotective;
KW	antitumor; hypocoenative; antidiabetic; antidiabetic; antidiabetic; muscular;
KW	antidiabetic; anorectic; gene therapy; cell therapy; antisense therapy;
KW	human; ss.
XX	XX
OS	Homo sapiens.
XX	XX
PH	Key
FT	CDS
FT	Location/Qualifiers
FT	/*tag= a
FT	/product= "cloaked-2 polypeptide"
FT	37..105
FT	/*tag= b
FT	mat_peptide
FT	106..759
FT	/*tag= c
FT	/note= "see ABB07207"
XX	XX
XX	WO200192308-A2.
XX	06-DEC-2001.
XX	XX

PE 29-MAY-2001; 2001WO-US17478.
 XX
 PR 01-JUN-2000; 2000US-208550P.
 PR 04-AUG-2000; 2000US-223542P.
 XX
 XX
 PA (AMGE-) AMGEN INC.
 XX
 XX
 PI Paszty CJ, Gao Y;
 XX
 DR WPI; 2002-114325/15.
 DR P-PSDB; ABB07207; ABB07209.
 XX
 XX
 PT New human and mouse cystine-knot polypeptide designated as Cloaked-2,
 PT for treating or preventing kidney, heart (e.g. myocardial infarction)
 PT or liver (e.g. hepatitis) diseases -
 XX
 PS Claim 1; Fig 1; 170pp; English.

The invention relates to polypeptides comprising a cysteine knot motif and designated as Cloaked-2, derived from human and mouse. The cloaked-2 polypeptides can be expressed by standard recombinant methodology. The cloaked-2 polynucleotides are useful in gene therapy and antisense therapy. The cloaked-2 polypeptides and polynucleotides are useful for treating, preventing, ameliorating or detecting diseases and disorders of the kidney (e.g. anaemia, hypertension or low blood pressure), heart (e.g. cardiac hypertrophy, congestive heart failure, myocardial infarction, arrhythmias, atherosclerosis, hypertension or low blood pressure), skeletal muscle (e.g. muscular dystrophy or cachexia), placenta (e.g. congenital abnormalities or miscarriage), liver (e.g. hepatitis or cirrhosis), pancreas (e.g. diabetes or pancreatitis), thyroid (e.g. Grave's disease or myxedema) or adrenal cortex (e.g. Cushing's disease or Addison's disease), homeostasis or metabolic diseases (e.g. obesity, cancer or myopathies), infections, or autoimmune diseases. Selective binding agents may be used to modulate the biological activities of cloaked-2 polypeptides or to detect cloaked-2 polypeptide levels in a sample. Transgenic non-human animals are useful for drug candidate screening. The present sequence represents a cDNA encoding the human cloaked-2 polypeptide.

Sequence 759 BP; 125 A; 282 C; 244 G; 108 T; 0 other;

Query Match	100.0%	Score 759;	DB 24;	length 759;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 759; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0

Qy	1	TATTTGAAGTGGCGTGGCTCTCTTTGGCTGTATCAATGAGAGCTCCCACTGGGCGTGTG	60
Db	1	TACTTGAAAGTGGCGTGGCTCTCTTGTGTGTATCAATGACGTCCCACTGGGCGTGTG	60
Qy	61	CTCGTCTGGCTCTGGTACACAGACCTTCTCGTGTAGTGAAGGCGCAGGGGTGTGACAGCG	120
Db	61	CTCGTCTGGCTCTGGTACACAGCTTCTCGTGTAGTGAAGGCGCAGGGGTGTGACAGCG	120
Qy	121	TTTCAGAAATGATGCCACGGAAATCATCCCGAGCTCGAGATACCCGAGCTTCCACCG	180
Db	121	TTTCAGAAATGATGCCACGAAATCATCCCGAGCTCGAGATACCCCGAGCTTCCACCG	180
Qy	181	GAGCTGAGAAACAACAAGACATGTAAACGGGGCGAGAAACGGAAGGGGGAGCTCCCAACAC	240
Db	181	GAGCTGAGAAACAACAAGACATGTAAACGGGGCGAGAAACGGAAGGGGGAGCTCCCAACAC	240
Qy	241	CCCTTTGAGACCAAAAGACGTGTCCAGATACAGCTTCCCGAGCTGCACCTTCAACCCGCTAC	300
Db	241	CCCTTTGAGACCAAAAGACGTGTCCAGATACAGCTTCCCGAGCTGCACCTTCAACCCGCTAC	300
Qy	301	GTTGACCGATGGGCGTGTCCGCAAGCGCAAGCGGTCACTGAGCTGTGTGCTTCCGGCCAG	360
Db	301	GTTGACCGATGGGCGTGTCCGCAAGCGCAAGCGGTCACTGAGCTGTGTGCTTCCGGCCAG	360
Qy	361	TGGGCGCCGGGCGGCTGTGCTCCCAACCGCATTCGGCGCGGCAAGTGGTGGCGACTAGT	420
Db	361	TGGGCGCCGGGCGGCTGTGCTCCCAACCGCATTCGGCGCGGCAAGTGGTGGCGACTAGT	420

OY	421	GGGCCGCACTTCGCTGATGCCGACCGCTACCGGAGGCGTGCAGCTGTGT	480
Db	421	GGGCGCGCACTTCGCTGCATCCCGACCGCTACCGGAGGCGTGCAGCTGTGT	480
OY	481	CCCGTGTGTGAAGGCGCGCGCGCAAGGTGTGTGTGCTGTGTCAAGTGCAG	540
Db	481	CCCGTGTGTGAAGGCGCGCGCGCAAGGTGTGTGTGCTGTGTCAAGTGCAG	540
OY	541	CGCCTCACCGCTTCCACMACAGTCGAGCTCAAGACTTCGAGACCGAGCGCTCG	600
Db	541	CGCCTCACCGCTTCCACMACAGTCGAGCTCAAGACTTCGAGACCGAGCGCTCG	600
OY	601	CCGCAAGAGGCGCGAGCGCGCGCGCCCGGAGCGCCAAAGCAACAGCGCGAG	660
Db	601	CCGCAAGAGGCGCGAGCGCGCGCGCCCGGAGCGCCAAAGCAACAGCGCGAG	660
OY	661	CTGGAGAAAGCTACTAAGCGCGCGCGCGCCCTCCCAACCGGCGGCGCGCGCT	720
Db	661	CTGGAGAAAGCTACTAAGCGCGCGCGCGCCCTCCCAACCGGCGGCGCGCGCT	720
OY	721	GAAACCGGCGCCACATTTCTGTCTGTGCGCGTGT	759
Db	721	GAAACCGGCGCCACATTTCTGTCTGTGCGCGTGT	759

RESULT 2

ID AAA29055 standard; cDNA; 2301 BP

AC AAA29055;

DT 12-SEP-2000 (first entry)

DE Human TGF-beta binding protein (BEER) cDNA

osteopathic; transforming growth factor-beta; TGF-beta; binding protein; KW

KW bone mineralization; ss.

OS Homo sapiens

aa	Key	Location/Qualifiers
FH		

CBS
 11
 ET
 10:00
 / * t a d = a

/product=101-beta_binning_product

PN WO2000032113-A1
XX

08-JUN-2000.
PD
YY

PF 24-NOV-1999; 99WU-US2/990-
XX

PR 27-NOV-1998; 98US-0110283.

PA (DARW-) DARWIN DISCOVERY LTD.

PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepke RW,

XX

DR P-PSDB; AAY96429.

PT Nucleic acids (I) encoding a transforming growth factor beta binding

PT osteoporosis and fractures

PS Claim 1; Page 114-115; 162pp; English.

This cDNA encodes a human transforming growth factor-beta (TGF-beta)

to the chromosome 17q12-21. The cDNA and protein may be used for

prevention, treatment and diagnosis of various inappropriate BEER expression. For example, they may be used to treat


```

QY 3 CTGGAAGTGGCGCTGCTCTCTGCTGTATCATGACGCTTCCAGCTGAGCTGTCT 62
DB 28 CTGGAAGTGGCGCTGCTCTCTGCTGTATCATGACGCTTCCAGCTGAGCTGTCT 87
QY 63 CGTCTGCTGTGTATCACACAGCTTCCGTGTATGAGGAGGCGAGGGGAGGCGCTT 122
DB 88 CGTCTGCTGTGTATCACACAGCTTCCGTGTATGAGGAGGCGAGGGGAGGCGCTT 147
QY 123 CAAGAAATGATGCCACGGAATATCATCCCGAGCTGAGAGTATCCCGAGCTTCCAGCGA 182
DB 148 CAAGAAATGATGCCACGGAATATCATCCCGAGCTGAGAGTATCCCGAGCTTCCAGCGA 207
QY 183 GCTGTGAGAACAAACAAGCACTGAACCGGGGCGGGAACGAGGGGCGCTTCCCAACACCC 242
DB 208 GCTGTGAGAACAAACAAGCACTGAACCGGGGCGGGAACGAGGGGCGCTTCCCAACACCC 267
QY 243 CTTTGAAGACAAAGAAGTGTCCGAGTACAGCTGCGGAGCTGACCTTCAACCGCTACGT 302
DB 268 CTTTGAAGACAAAGAAGTGTCCGAGTACAGCTGCGGAGCTGACCTTCAACCGCTACGT 327
QY 303 GACCGATGAGCGCGTGTGCGGACGCGCAAGCGGTACCGAGCTGTGTCTCCGCGCAAGTG 362
DB 328 GACCGATGAGCGCGTGTGCGGACGCGCAAGCGGTACCGAGCTGTGTCTCCGCGCAAGTG 387
QY 363 GCGCGCGCGCGCTGCTGCTGCTCCAAACCGCATCGCGCGCGGCAAGTGTGTGTGTG 422
DB 388 GCGCGCGCGCGCTGCTGCTGCTCCAAACCGCATCGCGCGCGGCAAGTGTGTGTGTG 447
QY 423 GCGCGATCTTCCGCTGATCTCCCGACCGCTACCGCGCGCGAGCGGTGTGTGTGTGT 482
DB 448 GCGCGATCTTCCGCTGATCTCCCGACCGCTACCGCGCGCGAGCGGTGTGTGTGTGT 507
QY 483 CGGTGTGAGAGCGCGCGCGCGCGCGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 542
DB 508 CGGTGTGAGAGCGCGCGCGCGCGCGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 567
QY 543 CTTTACCGCGCTTCCAAACAAGTGTGAGCTTCAAGGAGCTTGTGTGTGTGTGTGTGT 602
DB 568 CTTTACCGCGCTTCCAAACAAGTGTGAGCTTCAAGGAGCTTGTGTGTGTGTGTGTGT 627
QY 603 GCAGAAAGGCGCGAAGCGCGCGCGCGCGCGCGAGCGCGCAACGAGCGCGAGCT 662
DB 628 GCAGAAAGGCGCGAAGCGCGCGCGCGCGCGCGAGCGCGCAACGAGCGCGAGCT 687
QY 663 GGAAGACGCTTCTAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTGA 722
DB 688 GGAAGACGCTTCTAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTGA 747
QY 723 ACCGCGCGCGCGCATTTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 759
DB 748 ACCGCGCGCGCGCATTTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 784

```

RESULT 4
ABK69992
ID ABK69992 standard; DNA, 2329 BP.

AC ABK69992;
DT 15-JUL-2002 (first entry)
XX

DE cDNA encoding human Pro peptide #32.

XX Human; ss; gene; PRO; secreted protein; transmembrane protein;
KW genetic disorder; tumour; cancer.
XX

OS Homo sapiens.

XX W0200224888-A2.
XX

XX 28-MAR-2002.
XX

PF 29-AUG-2001; 2001WO-US27099.
XX
PR 01-SEP-2000; 2000US-229896P.
PR 05-SEP-2000; 2000US-230621P.
PR 22-SEP-2000; 2000US-235147P.
PR 10-NOV-2000; 2000WO-US30873.
PR 12-JAN-2001; 2001US-261878P.
PR 16-JAN-2001; 2001US-261910P.
PR 16-JAN-2001; 2001US-261938P.
PR 16-JAN-2001; 2001US-262150P.
PR 25-JAN-2001; 2001US-264395P.
PR 02-FEB-2001; 2001US-266421P.
PR 09-FEB-2001; 2001US-267623P.
PR 28-FEB-2001; 2001WO-US06520.
PR 09-MAR-2001; 2001US-274399P.
PR 03-APR-2001; 2001US-280982P.
PR 04-APR-2001; 2001US-282129P.
PR 04-APR-2001; 2001US-282199P.
PR 09-MAY-2001; 2001US-290589P.
PR 25-MAY-2001; 2001WO-US17092.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19682.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
XX
PA (GETH) GENENTECH INC.
XX
PI Baker KP, Baton DL, Filvaroff E, Goddard A, Grimaldi JC;
PI Gunney AL, Smith V, Stephan J, Watanabe CK, Wood WT, Zhang Z;
PI Fong S;
XX
DR MPI; 2002-362426/39.
XX
PS P-PSDB; ABG34061.
XX
PT New PRO polypeptides and polynucleotides encoding the polypeptides,
PT useful in gene therapy, chromosome identification, tissue typing, or
PT for genetic analysis of individuals with genetic disorders
XX
DR Claim 2; Figure 63; 218pp; English.
XX
PS This invention relates to the cDNA and protein sequences of novel
CC secreted and transmembrane polypeptides PRO polypeptides. The
CC invention also comprises a method for producing the proteins of the
CC invention by recombinant means and antibodies specific for the protein
CC of the invention. The antibody may be used for detecting the PRO
CC proteins of the invention and may be used to modify their activity.
CC polynucleotides may be used as hybridisation probes for a cDNA library
CC to isolate the full-length PRO cDNA or to isolate other cDNAs, to
CC construct hybridisation probes for mapping the gene which encodes that
CC PRO and for genetic analysis of individuals with genetic disorders, in
CC assays to identify other proteins or molecules involved in binding
CC reaction, to generate transgenic animals or knock-out animals which in
CC turn are useful in the development and screening of therapeutically
CC useful reagents, for chromosome identification, and tissue typing. The
CC PRO polypeptides are useful in gene therapy, and as molecular weight
CC markers for protein electrophoresis purposes. The sequences may
CC also be used to detect overexpression on PRO polypeptides in cancerous
CC tumours and for screening for differentially expressed genes using
CC microarray technology. The present sequence represents a cDNA encoding
CC a human PRO protein of the invention.
XX
SQ Sequence 2329 BP; 587 A; 645 C; 612 G; 485 T; 0 other;

Query Match 99.7%; Score 757; DB 24; Length 2329;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 757; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 3 CTGGAAGTGGCGCTGCTCTCTGCTGTATCATGACGCTTCCAGCTGAGCTGTCT 62
DB 28 CTGGAAGTGGCGCTGCTCTCTGCTGTATCATGACGCTTCCAGCTGAGCTGTCT 87
QY 63 CGTCTGCTGTGTATCACACAGCTTCCGTGTATGAGGAGGCGAGGGGAGGCGCTT 122

```


Db	361	AAACCCTCGGCTCCGGCCAAAGTGTGGCGACTGATGAGGCCCACTTCCGCTGCATCCCC	420
Qy	445	GACCCTACCGCGCGCAGAGCGGTGCACTGCTGTGTCGCCGTGATGAGCGCGCGCGG	504
Db	421	GACCGCTACCGCGCGCAGAGCGGTGCACTGCTGTGTCGCCGTGATGAGCGCGCGCGG	480
Qy	505	CGCAAGGTGCGCTGTGTGGCTCTGTGCATGTGCAGAGCGCTCACCCGCTTCCCAACCCAG	564
Db	481	CGCAAGGTGCGCTGTGTGGCTCTGTGCATGTGCAGAGCGCTCACCCGCTTCCCAACCCAG	540
Qy	565	TGCGAGCTCANGACTTGGGACCGAGGCGCGCTTGGGCGCGCAGAGAGGCGGAAAGCGCGG	624
Db	541	TGCGAGCTCANGACTTGGGACCGAGGCGCGCTTGGGCGCGCAGAGAGGCGCGAAGCGCGG	600
Qy	625	CCCCCGCGCCGAGAGCGCCAAAGCCCAACCGAGGCGCAGCTGAGGAAGCGCTACTAGAGCCG	684
Db	601	CCCCCGCGCCGAGAGCGCCAAAGCCCAACCGAGGCGCAGCTGAGGAAGCGCTACTAGAGCCG	660
Qy	685	CCCGGCGCCCTCCCAACCGAGGCGCGCCCGCGCCTGTAACCCGCGGCCCACTTCTGTC	744
Db	661	CCCGGCGCCCTCCCAACCGAGGCGCGCCCGCGCCTGTAACCCGCGGCCCACTTCTGTC	720
Qy	745	CTCTGCGCGTGGTTT	759
Db	721	CTCTGCGCGTGGTTT	735

RESULT 6
AAA29061
ID AAA29061 standard; cDNA; 2301 BP.

DT 12-SEP-2000 (first entry)

DE Mutant human TGF-beta binding protein (BEER) cDNA..

AA osteopathic; transforming growth factor-beta; TGF-beta; binding protein,
KM BEER; chromosome 17q12-21; gene therapy; antisense therapy; fracture;
KM bone mineralization; mutant; sclerosteosis; ss.

OS Homo sapiens.

Key	Location/Qualifiers
EH	48..119
FT	/*tag= a
FT	/note= "mutant BEER created by sclerosteosis
FT	nonsense mutation"

PN WO200032773-A1.

PD 08-JUN-2000.

PF 24-NOV-1999; 99WO-US27990

PR 27-NOV-1998; 98US-0110283

PA (DARW-) DARWIN DISCOVERY LTD.

PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JJ, Raepel SM,
Winkler DG.

XX
WPT 2000-412221/35

DR P-PSDB; AAY96435.
yy

protein useful for identifying agents for treating osteopenia, PT Nucleic acids (1) encoding a transforming growth factor beta

PT osteoporosis and fractures

PS Claim 2/; page 11-110; 10zpp; English
XY

CC This cDNA encodes a human chondrocyte growth factor

(TGF-beta) binding protein (BEP) produced as a result of a nonsense mutation in the BEP coding sequence (C to T mutation at position 117). This mutation has been linked to sclerosteosis in humans by linkage analysis of an Afrikaaner population in South Africa. The hBEP gene has been localized to the chromosome 17q12-21. The cDNA and protein may be used for prevention, treatment and diagnosis of diseases associated with inappropriate BEP expression. For example, they may be used to treat disorders associated with decreased TGF-beta B expression. The cDNA or vectors may be administered to treat patients with recidivating mutations or deletions in a patient's genome that affect the activity of BEP by expressing inactive proteins or to supplement the patients own production of BEP polypeptides. The nucleic acids may be used for recombinant production of BEP, gene therapy, antisense therapy, as probes for diagnostic assays and for functional studies. BEP may be used to raise antibodies and for identification of BEP modulators. BEP antagonists may be used to increase bone mineral content for the treatment of disorders such as osteopenia, osteoporosis, fractures and other disorders associated with low mineral content.

SQ Sequence 2301 BP; 568 A; 634 C; 614 G; 485 T; 0 other

Query Match	93.38%	Score 708;	DB 21;	Length 2501
	90.08%	Prod No. 0.		

Matches	758;	Conservative	0;	Mismatches	1;	Indels	0;	Gaps	0
---------	------	--------------	----	------------	----	--------	----	------	---

1 TACTGGAAGTGGCTGTCCTCTCTGCTGGTAGCATGACAGTCCCACTGGCCCTGTGT 60
 2 |||||
 3 |||||
 4 |||||
 5 |||||
 6 |||||
 7 |||||
 8 |||||
 9 |||||
 10 |||||
 11 |||||
 12 |||||
 13 |||||
 14 |||||
 15 |||||
 16 |||||
 17 |||||
 18 |||||
 19 |||||
 20 |||||
 21 |||||
 22 |||||
 23 |||||
 24 |||||
 25 |||||
 26 |||||
 27 |||||
 28 |||||
 29 |||||
 30 |||||
 31 |||||
 32 |||||
 33 |||||
 34 |||||
 35 |||||
 36 |||||
 37 |||||
 38 |||||
 39 |||||
 40 |||||
 41 |||||
 42 |||||
 43 |||||
 44 |||||
 45 |||||
 46 |||||
 47 |||||
 48 |||||
 49 |||||
 50 |||||
 51 |||||
 52 |||||
 53 |||||
 54 |||||
 55 |||||
 56 |||||
 57 |||||
 58 |||||
 59 |||||
 60 |||||
 61 |||||
 62 |||||
 63 |||||
 64 |||||
 65 |||||
 66 |||||
 67 |||||
 68 |||||
 69 |||||
 70 |||||
 71 |||||

120

131

...TCTCCGAGCTACCCGAGCCTCCACCG 180

191 CCGCAGGCTCCACCG 192

...GAGAAACCGAGGCCCTCCCAACCAC 240

100 CAGCTCAGTACCAACAATGAAACCGGGCCGAGAAGCGGAGCGGCCTCCCCACCAC 25

241 GCGGTTTCAACCAACAGCTGTCCGACTACAGCTGCCGCGAGCTGCACCTCACCCTAC 30

353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

301 CTGAGGCGGTCACCGAGCTGGTGCTCCGGCCAG 36

312 ATGAGGCTGACAGCCGCTGCTCCGCCAG 37

361 TACGCGCGCGCGCTGTGTCAGCCATCGGCCCGCGCAAGTGGTGCGACTTACT 42

272 TCAGGCGGCGCTGTGCCCAAGCCATCGCCCGGCAAGTGTGGCACCTAGT 43

431 ccgcccgacttcgcgcctgcatcccccgaaccgctaccgcgcgcgcagcgcgtgcagctgctgtgt 48

432 GGGCCGACTTCGGTGCATCCCCGACCGGCTACCGCGCGCAGGCGGTGCAGCTGCTGTGT 49

481 C C C G G T G T T A G A G C C C C C G C C G C A A G G T G C G C C T G T G G C C T C G T G C A A G T G C A A G 54

492 C C C G G T G T G A G C C C C C C C C C G C A A G G T G C G C C T G G T G C C T C G T G C A A G T G C A A G 55

541 CGCCTCACCCTTCCACACACCAAGTCGAGCTCAGGACTTCGGACCGAGGCCGCTCGG 60

552 CGCCTCACCCGCTTCCACACCAACAGTCGAGCTCAAGACTTCGGGACCGAGCCGCTCGG 61

601 CCGCAGAGGGCCGGAAGCCGGCCCCCGCCCGGAGCCCAAGCCCAACCAAGGCCGAG 66

Db 612 CCGCAGAGGGCCGGAAGCCGCGGCCCGCGCCCGGAGCGCCCAAGCCAAACAGGCCGAG 67

661 CTGAGACGCCTACTAGAGCCCGCCCGCCCTCCCCACCGCGGCGCCCGGCCCT 72

Db 672 CTGAGACGCCCTACTAGAGCCCGCCCGCCCTCCCCACCGCGCGCCCGCCCT 73

Qy 721 GAACCCGCCCCCAGATTCTGCTCTGCGCGTGTGTT 759
Db 732 GAACCCGCCCCCAGATTCTGCTCTGCGCGTGTGTT 770

RESULT 7
AAA29062

ID AAA29062 standard; cDNA; 2301 BP.

AC AAA29062;

DT 12-SEP-2000 (first entry)

XX Human TGF-beta binding protein (BEER) variant P38R cDNA.

XX osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
KW BEER; variant; P38R; gene therapy; antineoplastic therapy; fracture;
KM chromosome 17q12-21; bone mineralization; ss.

XX Homo sapiens.

FH Key Location/Qualifiers
FT CDS 48..689
FT /tag= a
FT /label= BEER_variant_P38R
FT /product= TGF-beta_binding_protein

XX MO200032773-A1.

PD 08-JUN-2000.

XX 24-NOV-1999; 99MO-US27990.

XX 27-NOV-1998; 98US-0110283.

XX (DARW-) DARWIN DISCOVERY LTD.

XX Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepker BW;
PI Van Ness J, Winkler DG;DR WPI; 2000-412321/35.
DR P-PSDB; AAY96436#.PT Nucleic acids (I) encoding a transforming growth factor beta binding
PT protein, useful for identifying agents for treating osteopenia,
PT osteoporosis and fractures

PS Claim 27; Page 120-121; 162pp; English.

XX This cDNA encodes a variant human transforming growth factor-beta
XX (TGF-beta) binding protein designated BEER P38R. The encoded protein
XX comprises a substitution of arginine for the wild-type proline at
XX residue 38. The cDNA and protein may be used for prevention, treatment
XX and diagnosis of diseases associated with inappropriate BEER expression.
XX For example, they may be used to treat disorders associated with
XX decreased TGF-beta BP expression. The cDNA or vectors may be administered
XX to treat diseases by rectifying mutations or deletions in a patient's
XX genome that affect the activity of BEER by expressing inactive proteins
XX or to supplement the patients own production of BEER polypeptides. The
XX nucleic acids may be used for recombinant production of BEER gene
XX therapy, antisense therapy, as probes for diagnostic assays and for
XX functional studies. BEER may be used to raise antibodies and for
XX identification of BEER modulators. BEER antagonists may be used to
XX increase bone mineral content for the treatment of disorders such as
XX osteopenia, osteoporosis, fractures and other disorders associated with
XX low mineral content.

SQ Sequence 2301 BP; 568 A; 634 C; 615 G; 484 T; 0 other;

Query Match 93.3%; Score 708; DB 21; Length 2301;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 758; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TACGGAAGTGGCGTGGCCCTCTGCTGCTGATGACATGACAGCTCCCACTGACCTGTGT 60
Db 12 TACTGGAAGTGGCGTGGCCCTCTGCTGCTGATGACATGACAGCTCCCACTGACCTGTGT 71
Qy 61 CTCGTCTGCTCTCTGTATACACAGCTTCCGTGTATGAGAGGCGCAGGGGTGAGGCG 120
Db 72 CTCGTCTGCTCTCTGTATACACAGCTTCCGTGTATGAGAGGCGCAGGGGTGAGGCG 131
Qy 121 TTCAAGAAATGATGACGAGGAATCATCCCGAGCTGGAGAGTACCCGAGGCTCACCG 180
Db 132 TTCAGAAATGATGACGAGGAATCATCCCGAGCTGGAGAGTACCCGAGGCTCACCG 191
Qy 181 GAGCTGAGAAACAAAGACCATGAACCGGGGAGAAACGAGGCGGCTTCCGACAC 240
Db 192 GAGCTGAGAAACAAAGACCATGAACCGGGGAGAAACGAGGCGGCTTCCGACAC 251
Qy 241 CCCTTTGAGAACCAAGACGTGTCCAGATGACAGTCCCGGAGCTCACTTCAACCGCTAC 300
Db 252 CCCTTTGAGAACCAAGACGTGTCCAGATGACAGTCCCGGAGCTCACTTCAACCGCTAC 311
Qy 301 GTGACCGATGGGCGCTGCGCGAGCGCAAGCGGTCAACGAGCTGTGCTCCGCGCAG 360
Db 312 GTGACCGATGGGCGCTGCGCGAGCGCAAGCGGTCAACGAGCTGTGCTCCGCGCAG 371
Qy 361 TCGGCGCCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db 372 TCGGCGCCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 431
Qy 421 GGGCGCGACTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db 432 GGGCGCGACTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 491
Qy 481 CCGGCTGTGAGAGCGCGCGCGCGCAAGGTGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 540
Db 492 CCGGCTGTGAGAGCGCGCGCGCGCAAGGTGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 551
Qy 541 CGCTTACCCGCTTCCACCAACGATGCGAGCTCAAGAGCTTGGAGACCGAGCGCTGCTGG 600
Db 552 CGCTTACCCGCTTCCACCAACGATGCGAGCTTCAAGAGCTTGGAGACCGAGCGCTGCTGG 611
Qy 601 CCGGAGAAAGGCGGAGACCG 660
Db 612 CCGGAGAAAGGCGGAGACCG 671
Qy 661 CTGAGAAAGGCTTCAAGAGCT 720
Db 672 CTGAGAAAGGCTTCAAGAGCT 731

RESULT 8

ID AAA29056 standard; cDNA; 2301 BP.

AC AAA29056;

DT 12-SEP-2000 (first entry)

XX Human TGF-beta binding protein (BEER) variant V10I cDNA.

XX osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
KW BEER; variant; V10I; gene therapy; antineoplastic therapy; fracture;
KM chromosome 17q12-21; bone mineralization; ss.

XX Homo sapiens.

FH Key Location/Qualifiers
FT CDS 48..689
FT /tag= a
FT /label= BEER_variant_V10I

FT /product= tgf-beta_binding_protein
 XX WO200032773-A1.
 PN 08-JUN-2000.
 PD XX
 PF 24-NOV-1999; 99WO-US27990.
 PR 27-NOV-1998; 98US-0110283.
 XX
 PA (DARW-) DARWIN DISCOVERY LTD.
 XX
 PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paepfer BW;
 PI Van Ness J, Winkler DG;
 DR WPI; 2000-412321/35.
 DR P-PSDB; AAY96430.
 XX
 PT Nucleic acids (1) encoding a transforming growth factor beta binding
 PT protein, useful for identifying agents for treating osteopenia,
 PT osteoporosis and fractures
 PS
 PS Claim 1; Page 118-119; 162pp; English.
 CC This cDNA encodes a variant human transforming growth factor-beta
 CC (TGF-beta) binding protein designated BEER VI01. The encoded protein
 CC comprises a substitution of isoleucine for the wild-type valine at
 CC residue 10. The cDNA and protein may be used for prevention, treatment
 CC and diagnosis of diseases associated with inappropriate BEER expression.
 CC For example, they may be used to treat disorders associated with
 CC decreased TGF-beta BP expression. The cDNA or vectors may be administered
 CC to treat diseases by rectifying mutations or deletions in a patient's
 CC genome that affect the activity of BEER by expressing inactive proteins
 CC or to supplement the patient's own production of BEER polypeptides. The
 CC nucleic acids may be used for recombinant production of BEER, gene
 CC therapy, antisense therapy, as probes for diagnostic assays and for
 CC functional studies. BEER may be used to raise antibodies and for
 CC identification of BEER modulators. BEER antagonists may be used to
 CC increase bone mineral content for the treatment of disorders such as
 CC osteopenia, osteoporosis, fractures and other disorders associated with
 CC low mineral content.
 CC
 CC Sequence 2301 BP; 569 A; 634 C; 614 G; 484 T; 0 other;
 SQ
 Query Match 86.6%; Score 657; DB 21; Length 2301;
 Best Local Similarity 99.7%; Pred. No. 1.6e-295;
 Matches 757; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 TACTGGAAGTGGCGTCCCTCTCTGAGTGTACATGACAGCTCCCACTGCGCTGTGT 60
 DB 12 TACTGGAAGTGGCGTCCCTCTCTGAGTGTACATGACAGCTCCCACTGCGCTGTGT 71
 QY 61 CTGATCTGCTGCTGTGTACACACAGCTTCCGTGTAGTGAAGGGCCAGGGGTGGCAGGG 120
 DB 72 CTGATCTGCTGCTGTGTACACACAGCTTCCGTGTAGTGAAGGGCCAGGGGTGGCAGGG 131
 QY 121 TTCAAGAAATGATGCGACGGAATCATCCCGAGCTCGGAGAGTACCCCGAGCTCCACCG 180
 DB 132 TTCAAGAAATGATGCGACGGAATCATCCCGAGCTCGGAGAGTACCCCGAGCTCCACCG 191
 QY 181 GAGCTGGAACAACAAGACCATGAAACCGGCGGAGAAACGAGGGGCGCTCCCAACAC 240
 DB 192 GAGCTGGAACAACAAGACCATGAAACCGGCGGAGAAACGAGGGGCGCTCCCAACAC 251
 QY 241 CCCTTTGAGACCAAGACGTGTCCGAGTACAGCTGCGCGAGCTGCACTTCAACCCGCTAC 300
 DB 252 CCCTTTGAGACCAAGACGTGTCCGAGTACAGCTGCGCGAGCTGCACTTCAACCCGCTAC 311
 QY 301 GTGACCATGGGCGGTGCGGAGCGCCAAAGCGGCTCAACCGAGTGTGTGCTCCGGCGAG 360
 DB 312 GTGACCATGGGCGGTGCGGAGCGCCAAAGCGGCTCAACCGAGTGTGTGCTCCGGCGAG 371
 QY 361 TGGCGCGCGGCGCTGTGCGCAACGCGCATGCGCGCGCAAGTGTGCGAGCTTGT 420

DB 372 TGGCGCGCGGCGCTGTGCGCAACGCCATCGGCGCGCAAGTGTGCGAGCTTGT 431
 QY 421 GGGCCCGACCTTCCGCTGATTCGCCGACCTACCGCGGAGGCGGTGCGAGCTGTGT 480
 DB 432 GGGCCCGACCTTCCGCTGATTCGCCGACCTACCGCGGAGGCGGTGCGAGCTGTGT 491
 QY 481 CCGGTGTGTGAGGCGCGCGCGCGCGCAAGGATGCGCTGTGTGCTGTGCAAGTGAAG 540
 DB 492 CCGGTGTGTGAGGCGCGCGCGCGCGCAAGGATGCGCTGTGTGCTGTGCAAGTGAAG 551
 QY 541 CGCTTCAACCGCTTCCCAACCAAGTGTGAGCTCAAGACTTTGGGACCGAGCGCGCTCG 600
 DB 552 CGCTTCAACCGCTTCCCAACCAAGTGTGAGCTCAAGACTTTGGGACCGAGCGCGCTCG 611
 QY 601 CCGCAAGAGGCGCGGAGCGCGCGCGCGCGCGCGCGCGCAAGGCGCAAGCGCAAGCGCGAG 660
 DB 612 CCGCAAGAGGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCAAGGCGCAAGCGCGAG 671
 QY 661 CTGGAAGACGCTACTAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 720
 DB 672 CTGGAAGACGCTACTAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 731
 QY 721 GAACCGCT 759
 DB 732 GAACCGCT 770
 RESULT 9
 AAA94051
 ID AAA94051 standard; DNA; 642 BP.
 XX
 XX AAA94051;
 AC
 XX 15-JAN-2001 (first entry)
 DT
 XX
 DE Human DAN/Cerberus-related protein 6 (hDCR6) cDNA exons 1 and 4.
 XX
 KW Human; DAN/Cerberus-related protein 6; hDCR6; morphogenic protein;
 KW antagonist; BMP; cell growth; cell differentiation; bone formation;
 KW gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 XX
 FH Key Location/Qualifiers
 FH CDS 1..642
 FT /*tag= a
 FT /product= "hDCR6 #2"
 FT /partial
 PN WO20005193-A2.
 XX
 PD 21-SEP-2000.
 XX
 PF 02-MAR-2000; 2000WO-US05537.
 XX
 PR 12-MAR-1999; 99US-0124118.
 XX
 PA (REGG-) REGENERON PHARM INC.
 XX
 PI Economides AN;
 XX
 DR WPI; 2000-638179/61.
 DR P-PSDB; AAB26106.
 XX
 PT Novel isolated, human DNA/Cerberus related protein 6 which include
 PT natural homologue, and polypeptides comprising DCR6 domain and nucleic
 PT acids encoding the proteins which are useful as probes and primers -
 XX
 PS Claim 2; Fig 3; 40pp; English.
 CC The present sequence comprises exons 1 and 4 of the human
 CC DAN/Cerberus-related protein 6 (hDCR6) coding sequence. It was isolated

	RESULT 6	
Db	721 CTTGCGCCTGGTGT 735	
LOCUS	AF326742	642 bp mRNA linear PRI 28-FEB-2001
DEFINITION	Cercopithecus aethiops sclerostin cDNA, complete cds.	
ACCESSION	AF326742	
VERSION	AF326742.1 GI:13161028	
KEYWORDS		
SOURCE		
ORGANISM	Cercopithecus aethiops. Cercopithecus aethiops Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae; Cercopithecinae; Cercopithecus.	
REFERENCE	1 (bases 1 to 642) Brunkow,M.E., Gardner,J.C., Van Ness,J., Paeppe,B.W., Kovacevich,B.R., Prolli,S., Skonier,J.E., Zhao,L., Sabo,P.J., Fu,Y.H., Alisch,R.S., Gillett,L., Colbert,T., Tacconi,P., Galae,D., Hamerema,H., Beighton,P. and Mulligan,J.T. Bone dysplasia scleroosteosis results from loss of the SOST gene product, a novel cystine knot-containing protein Am. J. Hum. Genet. 68 (3), 577-589 (2001)	
TITLE	2 (bases 1 to 642) JOURNAL MEDLINE PUBMED 21090529 1179006	
AUTHORS	Brunkow,M.E., Gardner,J.C., Van Ness,J., Paeppe,B.W., Kovacevich,B.R., Prolli,S., Skonier,J.E., Zhao,L., Sabo,P.J., Fu,Y.H., Alisch,R.S., Gillett,L., Colbert,T., Tacconi,P., Galae,D., Hamerema,H., Beighton,P. and Mulligan,J.T. Direct Submission Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631 22oth St. SE, Bothell, WA 98021, USA	
FEATURES	Location/Qualifiers source 1..642 /organism="Cercopithecus aethiops" /db_xref="taxon:9534" 1..642 /codon_start=1 /product="sclerostin" /protein_id="AAK13457.1" /db_xref="GI:13161029" /translation="MQPLALCLVLAFAFVEQGNQAIFNDATETIIPELGEYR EPPELENNKTMMNAENAGRGPRPHPEFKDSEVSCHEHTRTVTMPGRSAKPTE LVSGGCGGPAPLAIENATIGRWKMPSPDPDRICIDRRRAORVOLTCGGAAPRARAKVR LVASCKRKRLTRFNOSBLKDGFEDARPOGRKRPBPARGAANKAQALENAY"	
BASE COUNT	113 a 228 c 210 g 91 t	
ORIGIN		
Query Match	80.8%; Score 613.2; DB 9; Length 642;	
Best Local Similarity	97.2%; Pred. No. 3e-84;	
Matches 624; Conservative	0; Mismatches 18; Indels 0; Gaps 0;	
QY	37 ATGCAGCTCCCACTGCGCCCTGTGTCGTCTGCGCTGCTGTAACAACAACCCTTCGGTGTA 96	
Db	1 ATGCAGCTCCCACTGCGCCCTGTGTCGTCTGCGCTGCTGTAACAACAACCCTTCGGTGTA 60	
QY	97 GTGAGAGGCCCAAGGGGTGGGAGGCGTTCAAGAATGATGCCACGGAATCATCTCCGAGCTC 156	
Db	61 GTGAGAGGCCCAAGGGGTGGGAGGCGTTCAAGAATGATGCCACGGAATCATCTCCGAGCTC 120	
QY	157 GGAGAGTACCCCGAGGCTTCACCCGAGGTGGAGAGTGGAGAACCAAGAACAATGACACCGGCGGAG 216	
Db	121 GGAGAGTACCCCGAGGCTTCACCCGAGGTGGAGAGTGGAGAACCAAGAACAATGACACCGGCGGAG 180	
QY	217 AACGAGGGGCGGCGCTTCACCAACACCCCTTTTGAGACCAAGAAGCGTGTCCGAGTAGACGTGC 276	
Db	181 AATGAGAGGGCGGCGCTTCACCAACACCCCTTTTGAGACCAAGAAGCGTGTCCGAGTAGACGTGC 240	
QY	277 CGCAGAGTGCACCTTACCCCCGCTAGTAGTACCGATGGGCGCTGCGCGAGCGCACAGCGCGCTC 336	
Db	241 CGAAGCTGCACCTTACCCCCGCTAGTAGTACCGATGGGCGCTGCGCGAGCGCACAGCGCGCTC 300	

OY	337	ACCGAGCTGATGTCTCCGGCCAGTGGCCGCCGCCTCTGTGCCCCAAGCCTATCGGC	396
Db	301	ACCGATTGGTGTGTCTCCGGCCAGTGGCCGCCGCCTCTGTGCCCCAAGCCTATCGGC	360
OY	397	CSCGGCAAGTGTTGTCGAGCACTTAGTGGGCGCCGACTTCCGCTGCATCCCAGCGCTACGC	456
Db	361	CSCGGCAAGTGTTGTCGAGCACTTAGTGGGCGCCGACTTCCGCTGCATCCCAGCGCTACGC	420
OY	457	GCGCAGCGCGTGCACCTGTGTGTCTCCGGTGTGAAGCGCGCGCGCGCAAGTGGCGC	516
Db	421	GCGCAGCGCGTGCACCTGTGTGTCTCCGGTGTGAAGCGCGCGCGCGCAAGTGGCGC	480
OY	517	CTGGTGGCTCTGTGAAGTGCAGAGGGCCCTCACCCGCTTCCAACAACATCGAGAAGCTTAAG	576
Db	481	CTGGTGGCTCTGTGAAGTGCAGAGGGCCCTCACCCGCTTCCAACAACATCGAGAAGCTTAAG	540
OY	577	GACTTCGGGAAACCGAGGCGCTCTGCGCCGCGAGAGAGGCGCGAAGCGCGCGCGCGCGG	636
Db	541	GACTTCGGTCCCGAAGGCGCTCTGCGCCGCGAGAGAGGCGCGAAGCGCGCGCGCGCGG	600
OY	637	AGCGCCAAAGCCACACGAGCGCGAGCTGGAGAAAGCGCTTACTAG	678
Db	601	GCGGCGCAAAGCCATCAGGCGAGCTGGAGAAAGCGCTTACTAG	642
RESULT 7			
LOCUS	AF326741	674 bp	mRNA linear ROD 28-FEB-2001
DEFINITION	Rattus norvegicus sclerostin mRNA, complete cds.		
ACCSSION	AF326741		
VERSION	AF326741.1	GI:13161025	
KEYWORDS			
SOURCE	Rattus norvegicus.		
ORGANISM	Rattus norvegicus.		
REFERENCE			
AUTHORS	1 (bases 1 to 674)		
TITLE	Brunkow,M.E., Gardner,J.C., Van Ness,J., Pieper,B.W., Kovacevich,B.R., Prohl,S., Skonier,J.E., Zhao,L., Sabo,P.J., Fu,Y.H., Alishch,R.S., Gillet,L., Colbert,T., Tacconi,P., Galas,D., Hamersma,H., Beighon,P. and Mulligan,J.T. Bone dysplasia sclerosteosis results from loss of the SOST gene product, a novel cystine knot-containing protein Am. J. Hum. Genet. 68 (3), 577-589 (2001)		
JOURNAL			
MEDLINE			
PUBMED	21090529		
REFERENCE	2 (bases 1 to 674)		
AUTHORS	Brunkow,M.E., Gardner,J.C., Van Ness,J., Pieper,B.W., Kovacevich,B.R., Prohl,S., Skonier,J.E., Zhao,L., Sabo,P.J., Fu,Y.H., Alishch,R.S., Gillet,L., Colbert,T., Tacconi,P., Galas,D., Hamersma,H., Beighon,P. and Mulligan,J.T. Direct Submmission Submitted (07-DEC-2000) Genomics, Cellect Chiroscience Inc., 1631 220th St. SE, Bothell, WA 98021, USA		
TITLE			
JOURNAL			
FEATURES			
SOURCE			
1..674	/organism="Rattus norvegicus"		
	/strain="Sprague-Dawley"		
	/db_xref="taxon:10116"		
1..32			
33..674	/codon_start=1		
	/product="sclerostin"		
	/protein_id="AK13456.1"		
	/db_xref="GI:13161026"		
	/translation="MOI:SLAPCLAACLVAFAVAVESQMAFKNDAITELPGLREYR EPPOELFNNOTMARAENGSRPHHPDPTKOVSEYSCELYHTPEVTDPGRSAKPVEB LVSGOGEPAKLIPNALIGVKWMPENBPDRCLIDRYRARVDLLCPGAAPRSKVR LVASCKCKRLTRRHNSQLMDFPETARPKDKRPFRARGAKRANOAELENAV"		
BASE COUNT	128 a	240 c	208 g 98 t
ORIGIN			

Kamat, A., Karatas, A., Kelle, C., Lacroque, K., Lamazares, R., Landers, T., Lenoczky, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneses, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, J., Retter, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schniback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vasilev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zaitoun, J., Zembek, L., Zimmer, A. and Zody, M.

Whitehead Institute/MIT Center for Genome Research
Submitted (22-FEB-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 22, 2002 this sequence version RepeatMasker: g1.15055348.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
Genome Center
Web site: http://www-seg.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: L7949
Center clone name: 209_M_4

Only the last 80.1 kilobases of this clone are being submitted.
The remainder overlaps accession number AC004149 [WICR project L302].

FEATURES

source

Location/Qualifiers
1. 80117
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="RP11-209M4"
/clone_lib="RP11-209M4 Human Male BAC"
complement(146.434)
/rpt_family="AluSg"
complement(438.655)
/rpt_family="Mir"
complement(974.1129)
/rpt_family="MER20"
complement(1270.1438)
/rpt_family="AluJo"
1439.1465
/rpt_family="TTTAA)n"
complement(146.1584)
/rpt_family="AluJo"
complement(1785.1933)
/rpt_family="MERSB"
2339.2487
/rpt_family="MIR"
complement(2492.2787)
/rpt_family="AluSg"
complement(2796.2902)
/rpt_family="MIR"
complement(2903.3165)
/rpt_family="AluSg"
complement(3166.3193)
/rpt_family="MIR"
3572.3644
/rpt_family="MIR"
4524.4593
/rpt_family="MIR3"
complement(4755.4962)
/rpt_family="MIR"
6397.6738

repeat_region /rpt_family="L2"
complement(6766.6875)
/rpt_family="MLT1H"
repeat_region 7202.7270
/rpt_family="L2"
repeat_region 7739.7825
/rpt_family="L2"
complement(7826.8122)
/rpt_family="AluSc"
8123.8268
/rpt_family="L1MC4"
8269.8568
/rpt_family="AluJo"
8569.8631
/rpt_family="L1MC4"
complement(9151.9305)
/rpt_family="MIR"
9365.9390
/rpt_family="AT_rich"
complement(9405.9708)
/rpt_family="AluY"
complement(9709.9969)
/rpt_family="AluJo"
10197.10458
/rpt_family="MLT1K"
10758.10895
/rpt_family="MIR"
12676.12773
/rpt_family="MIR"
complement(12784.12997)
/rpt_family="MIR"
complement(13220.13513)
/rpt_family="AluJo"
13627.13787
/rpt_family="L2"
complement(13821.13953)
/rpt_family="MIR"
14746.14899
/rpt_family="MIR"
complement(14947.15144)
/rpt_family="MIR"
15520.15668
/rpt_family="AluJo"
15669.15960
/rpt_family="AluSx"
15961.16128
/rpt_family="AluJo"
16208.16400
/rpt_family="MIR"
16454.16586
/rpt_family="L2"
16707.16783
/rpt_family="MIR"
16852.16909

Query Match 68.1%; Score 516.8; DB 9; Length 80117;
Best Local Similarity 96.0%; Pred. No. 5e-70;
Matches 530; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 208 CCGGCGGAGAGAGGCGGCGCTCCCAACACCCCTTTAGACCAAGAGTGTCCGAG 267
DB 61263 CAGGGGTGGCCAGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 63204
QY 268 TACAGCTGCGGAGCTGCTACCTTACCGGCTACGTCAGTCCGATGGGCGGCGGCGGCGG 327
DB 61203 TACAGCTGCGGAGCTGCTACCTTACCGGCTACGTCAGTCCGATGGGCGGCGGCGGCGG 63144
QY 328 AAGCCGTCACCGAGCTGCTGCTCGGCGGAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 387
DB 61143 AAGCCGTCACCGAGCTGCTGCTCGGCGGAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 63084
QY 388 GCCATGCGCGGCGGAGTGGTGGGAGCTTGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 447


```

/rpt_family="MIR"
repeat_region 20398..20769
/rpt_family="LINE"
repeat_region 20893..21190
/rpt_family="AluSg"
complement(21275..21473)
/rpt_family="AluSg"
complement(21485..21556)
/rpt_family="(CA)n"
repeat_region 21567..21906
/rpt_family="MER2"
repeat_region 21907..21952
/rpt_family="MLT2CB"
repeat_region 21974..22115
/rpt_family="MIR"
repeat_region 22238..22540
/rpt_family="AluSx"
repeat_region 22577..22633
/rpt_family="LINE2"
complement(23190..23270)
/rpt_family="LINE2"
repeat_region 24194..24324
/rpt_family="FLAM_A"
repeat_region 24329..24643
/rpt_family="AluSp"
repeat_region 25542..25698
/rpt_family="LINE2"
repeat_region 25805..26001
/rpt_family="MIR"
repeat_region 26060..26174
/rpt_family="LINE2"
repeat_region 26343..26649
/rpt_family="AluSx"
complement(27491..27783)
/rpt_family="AluSg"
repeat_region 27862..27906
/rpt_family="LINE2"
repeat_region 28068..28114
/rpt_family="MIR"
repeat_region 28371..28652
/rpt_family="AluSg"
complement(29498..29604)
/rpt_family="GC-rich"
repeat_region 30074..30181
/rpt_family="MIR"
complement(30181..30249)
/rpt_family="LINE2"
repeat_region 30444..30517
/rpt_family="MIR"
complement(31852..32135)
/rpt_family="AluSg"
repeat_region 32675..33003
/rpt_family="MLT1A1"
complement(33042..33082)
/rpt_family="LINE2"
repeat_region 33363..33459
/rpt_family="MER5B"
complement(35475..35524)
/rpt_family="MLT2D"
repeat_region 35495..35591
/rpt_family="(GA)n"
complement(35561..35673)
/rpt_family="MLT2D"
repeat_region 35678..35935
/rpt_family="MER6B"
complement(35936..36233)
/rpt_family="AluSg"
complement(36236..36449)
/rpt_family="MER6B"
complement(36450..36867)
/rpt_family="MLT2D"

```

```

repeat_region complement(36876..37460)
/rpt_family="L1MD2"
repeat_region complement(37477..37777)
/rpt_family="AluSx"
repeat_region complement(37804..37948)
/rpt_family="L1"
repeat_region 37948..38130
/rpt_family="L1MD1"
repeat_region 38175..38253
/rpt_family="LINE2"

```

Query Match 68.1%; Score 516.8; DB 9; Length 94752;
 Best Local Similarity 96.0%; Pred. No. 4,8e-70;
 Matches 530; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

```

QY 208 CGGGGCGAGAAACGAGGAGGCGGCTCCACACACCCCTTTGAGACCAAGAAGTGTCCGAG 267
DB 6340 CAGGGGTTGAGCCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 6281
QY 268 TACAGCTGCGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 327
DB 6280 TACAGCTGCGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6221
QY 328 AAGCGGTTACAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 387
DB 6220 AAGCGGTTACAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6161
QY 388 GCCATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 447
DB 6160 GCCATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 6101
QY 448 CGCTACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 507
DB 6100 CGCTACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 6041
QY 508 AAGTGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 567
DB 6040 AAGTGTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5991
QY 568 GAGCTCAAGAGCTTGGGAGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 627
DB 5980 GAGCTCAAGAGCTTGGGAGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 5921
QY 628 CGGCGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 687
DB 5920 CGGCGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 5861
QY 688 GGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 747
DB 5860 GGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 5801
QY 748 TGGCGGTGGTTT 759
DB 5800 TGGCGGTGGTTT 5789

```

RESULT 12
 AF326740 638 bp mRNA linear ROD 28-FEB-2001
 LOCUS AF326740
 DEFINITION Mus musculus sclerostin mRNA, complete cds.
 ACCESSION AF326740
 VERSION AF326740.1 GI:13161022
 KEYWORDS

SOURCE

Mus musculus.
 Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 638)
 Brunkow, M.E., Gardner, J.C., Van Ness, J., Paepker, B.W.,
 Kovacevich, B.R., Prohl, S., Skonier, J.E., Zhao, L., Sabo, P.J.,
 Fu, Y.H., Altsch, R.S., Gillet, L., Colbert, T., Taconi, P., Galas, D.,
 Hamersma, H., Beignon, P. and Mulligan, J.T.
 Bone dysplasia sclerosteosis results from loss of the SOST gene

QY 637 AGCGCCAAAGCCACGAGCGGCTGAGAAAGCCTACTAG 678
 DB 595 GGAGCCAAAGCCACGAGCGGCTGAGAAAGCCTACTAG 636
 RESULT 14
 AF326738 532 bp mRNA linear MAM 28-FEB-2001
 LOCUS Bos taurus sclerostin mRNA, partial cds.
 DEFINITION AF326738
 ACCESSION AF326738.1 GI:13161016
 VERSION
 KEYWORDS
 SOURCE Bos taurus.
 ORGANISM Bos taurus.
 Bos taurus
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovidae; Bovinae; Bos.
 REFERENCE 1 (bases 1 to 532)
 AUTHORS Brunkow,M.E., Gardner,J.C., Van Ness,J., Paepfer,B.W.,
 Kovacevich,B.R., Prohl,S., Skonier,J.E., Zhao,L., Sabo,P.J.,
 Fu,Y.H., Alisch,R.S., Gillett,L., Colbert,T., Tacconi,P., Galas,D.,
 Hamersma,H., Beighton,P. and Mulligan,J.T.
 TITLE Bone dysplasia sclerostosis results from loss of the SOST gene
 product, a novel cysteine knot-containing protein
 JOURNAL Am. J. Hum. Genet. 68 (3), 577-589 (2001)
 MEDLINE 21090529
 PUBMED 11179006
 REFERENCE 2 (bases 1 to 532)
 AUTHORS Brunkow,M.E., Gardner,J.C., Van Ness,J., Paepfer,B.W.,
 Kovacevich,B.R., Prohl,S., Skonier,J.E., Zhao,L., Sabo,P.J.,
 Fu,Y.H., Alisch,R.S., Gillett,L., Colbert,T., Tacconi,P., Galas,D.,
 Hamersma,H., Beighton,P. and Mulligan,J.T.
 TITLE Direct Submission
 JOURNAL Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631
 220th St. SE, Bothell, WA 98021, USA
 FEATURES
 source location/Qualifiers
 1..532
 /organism="Bos taurus"
 /db_xref="taxon:9913"
 <1..>532
 /product="sclerostin"
 /codon_start=3
 /protein_id="AAK13453.1"
 /db_xref="GI:13161017"
 /translation="NDATRTIPRLGYPEPPLINKTKNRAENGRRPHHPETKDA
 SEYSRELAHFRVYDPCRSAPVTELVCSQCGARLPLVAIGRGKMKRSGDFR
 CIPDRYARQVQLCPGGAAPARKRVIAVSCCKRLTFHNSLKKDGFPAARPT
 GRKLRPRAGTKASRA"
 CDS
 BASE COUNT 95 a 199 c 176 g 62 t
 ORIGIN
 Query Match 61.6%; Score 467.8; DB 4; Length 532;
 Best Local Similarity 93.5%; Pred. No. 4.3e-62;
 Matches 500; Conservative 0; Mismatches 32; Indels 3; Gaps 1;
 QY 125 AGATGATGCGGAGGATATCCCGAGCTCGGAGGTAACCCGAGCTCCACCGGAGC 184
 DB 1 AGATGATGCGGAGGATATCCCGAGCTCGGAGGTAACCCGAGCTCCACCGGAGC 60
 QY 185 TGGAGAACCAACAGACCATGAAACCGGCGGAGAAACGAGGCGGCTCCACCAACCCCT 244
 DB 61 T--GAACACAAAGACCATGAAACCGGCGGAGAAACGAGGAGACCTCCACCAACCCCT 117
 QY 245 TTGAGACCAAGAGCGTGTCCGAGTACAGCTCCCGGAGCTGCACTTCAACCCGCTAAGTGA 304
 DB 118 TTGAGACCAAGAGCGTGTCCGAGTACAGCTCCCGGAGCTGCACTTCAACCCGCTAAGTGA 177
 QY 305 CCGAGTGGCGGTGGCGGAGCGGCAAGCGGATCACCGAGCTGTGTCTCCGGCCAGTGGC 364
 DB 178 CCGAGTGGCGGTGGCGGAGCGGCAAGCGGATCACCGAGCTGTGTCTCCGGCCAGTGGC 237
 QY 365 GCCGGCGCGCTGCTGCCAAGCGCATCGGCGGCGCAAGTGTGGCGACTAAGTGGC 424

DB 238 GCCCGGCGGCGCTGTCGCCAAGCCATCGGCGGCAAGTGTGGCGGCCCAAGCGGCGC 297
 QY 425 CCGACTTCGCTGATTCGCCGACCGCTACCGCGGAGCGGTGAGCTGTGTCCG 484
 DB 298 CCGACTTCGCTGATTCGCCGACCGCTACCGCGGAGCGGTGAGCTGTGTCTG 357
 QY 485 GTGTGAGGCGCGCGGCGGCGCAAGTGCAGCTGTGTGCTGTGTCAGAGTGAACGCGC 544
 DB 358 GCGGCGGCGCGCGCGCGCGCGGAGGTGTGCTGTGCTGTGTCAGAGTGAACGCGC 417
 QY 545 TCACCGCTTCCAAACAGTGTGAGCTCAAGGACTTTCGGAGCCGAGCGCTCGGCGCC 604
 DB 418 TCACCTGCTTCCAAACAGTGTGAGCTCAAGGACTTTCGGAGCGCGGCGCGCGCGC 477
 QY 605 AGAAGGCGCGGAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 659
 DB 478 AAGCGGCGCGGAGCTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 532
 RESULT 15
 AC073954 177744 bp DNA linear HTG 17-NOV-2000
 LOCUS Homo sapiens chromosome 17 clone RP11-147C10, WORKING DRAFT
 DEFINITION AC073954
 ACCESSION AC073954.3 GI:10988993
 VERSION AC073954
 KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 177744)
 AUTHORS Waterston,R.H.
 TITLE The sequence of Homo sapiens clone
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 177744)
 AUTHORS Waterston,R.H.
 TITLE Direct Submission
 JOURNAL Submitted (07-JUL-2000) Genome Sequencing Center, Washington
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,
 MO 63108, USA
 COMMENT On Oct 25, 2000 this sequence version replaced gi:1986849.
 ----- Genome Center -----
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc/index.shtml
 Project Information -----
 Center project name: H_NH014/C10
 ----- Summary Statistics -----
 Sequencing vector: M13; 100%
 Chemistry: Dye-terminator Big Dye; 0% of reads
 Chemistry: Dye-terminator Big Dye; 0% of reads
 Assembly program: Phrap; version 0.99019
 Consensus quality: 155679 bases at least Q40
 Consensus quality: 163329 bases at least Q30
 Consensus quality: 166780 bases at least Q20
 Insert size: 185000; agarose-fp
 Insert size: 183559; sum-of-contigs
 Quality coverage: 3.29 in Q20 bases; sum-of-contigs
 Quality coverage: 3.43 in Q20 bases; sum-of-contigs
 ----- NOTE: This is a 'working draft' sequence. It currently
 * consists of 34 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 1247: contig of 1247 bp in length
 * 1248 1347: gap of unknown length

```

* 1348 2762: contig of 1415 bp in length
* 2763 2862: gap of unknown length
* 2863 4773: contig of 1911 bp in length
* 4774 4873: gap of unknown length
* 4874 6326: contig of 1453 bp in length
* 6327 6426: gap of unknown length
* 6427 8873: contig of 2447 bp in length
* 8874 8974 10783: contig of 1810 bp in length
* 10784 10883: gap of unknown length
* 10884 12763: contig of 1880 bp in length
* 12764 12863: gap of unknown length
* 12864 15381: contig of 2518 bp in length
* 15382 15481: gap of unknown length
* 15482 17365: contig of 1884 bp in length
* 17366 17465: gap of unknown length
* 17466 19992: contig of 2527 bp in length
* 19993 20092: gap of unknown length
* 20093 23512: contig of 3420 bp in length
* 23513 25699: contig of 2087 bp in length
* 25700 25799: gap of unknown length
* 25800 28333: contig of 2534 bp in length
* 28334 32027: gap of unknown length
* 32028 32127: gap of unknown length
* 32128 36115: contig of 3988 bp in length
* 36116 36215: gap of unknown length
* 36216 39882: contig of 3667 bp in length
* 39883 39983: gap of unknown length
* 39984 44009: contig of 4027 bp in length
* 44010 47829: contig of 3720 bp in length
* 47830 47930 53768: contig of 5839 bp in length
* 53769 53868: gap of unknown length
* 53869 60215: contig of 6347 bp in length
* 60216 60315: gap of unknown length
* 60316 66477: contig of 6162 bp in length
* 66478 66577: gap of unknown length
* 66578 71476: contig of 4899 bp in length
* 71477 71576: gap of unknown length
* 71577 78883: contig of 7307 bp in length
* 78884 78983: gap of unknown length
* 78984 85952: contig of 6969 bp in length
* 85953 86053 93510: gap of unknown length
* 86054 93511 93610: gap of unknown length
* 93611 100718: contig of 7108 bp in length
* 100719 100818: gap of unknown length
* 100819 109795: contig of 8977 bp in length
* 109796 109895: gap of unknown length
* 109896 117790: contig of 7895 bp in length
* 117791 117890: gap of unknown length
* 117891 126044: contig of 8154 bp in length
* 126045 126144: gap of unknown length
* 126145 134442: contig of 8298 bp in length
* 134443 134542: gap of unknown length
* 134543 145663: contig of 11121 bp in length
* 145664 145763: gap of unknown length
* 145764 157189: contig of 11426 bp in length
* 157190 157289: gap of unknown length
* 157290 166822: contig of 9533 bp in length
* 166823 177744: gap of unknown length
* 166923 177744: contig of 10822 bp in length.

```

FEATURES

```

source
1.177744
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/clone="RP11-147C10"
1..1247
/notes="assembly_name:Contig26"
misc_feature

```

```

misc_feature 1348..2762 /note="assembly_name:Contig31"
misc_feature 2863..4773 /note="assembly_name:Contig34"
misc_feature 4874..6326 /note="assembly_name:Contig35"
misc_feature 6427..8873 /note="assembly_name:Contig36"
misc_feature 8974..10783 /note="assembly_name:Contig37"
misc_feature 10884..12763 /note="assembly_name:Contig38"
misc_feature 12864..15381 /note="assembly_name:Contig39"
misc_feature 15482..17365 /note="assembly_name:Contig40"
misc_feature 17466..19992 /note="assembly_name:Contig41"
misc_feature 20093..23512 /note="assembly_name:Contig42"
misc_feature 23613..25699 /note="assembly_name:Contig43"
misc_feature 25800..28333 /note="assembly_name:Contig44"
misc_feature 28434..32027 /note="assembly_name:Contig45"
misc_feature 32128..36115 /note="assembly_name:Contig46"
misc_feature 36216..39882 /note="assembly_name:Contig47"
misc_feature 39983..44009 /note="assembly_name:Contig48"
misc_feature 44110..47829 /note="assembly_name:Contig49"
misc_feature 47930..53768 /note="assembly_name:Contig50"
misc_feature 53869..60215 /note="assembly_name:Contig51"
misc_feature 60316..66477 /note="assembly_name:Contig52"
misc_feature 66578..71476 /note="assembly_name:Contig53"
misc_feature 71577..78883 /note="assembly_name:Contig54"
misc_feature 78984..85952 /note="assembly_name:Contig55"
misc_feature 86053..93510 /note="assembly_name:Contig56"
misc_feature 93611..100718 /note="assembly_name:Contig57"
misc_feature 100819..109795 /note="assembly_name:Contig58"
misc_feature 109896..117790 /note="assembly_name:Contig59"
misc_feature 117891..126044 /note="assembly_name:Contig60"
misc_feature 126145..134442 /note="assembly_name:Contig61"
misc_feature 134543..145663 /note="assembly_name:Contig62"
misc_feature 145764..157189 /note="assembly_name:Contig63"
misc_feature 157290..166822 /note="assembly_name:Contig64"
misc_feature 166923..177744 /note="assembly_name:Contig65"
misc_feature 177744..189888 /note="assembly_name:Contig66"

```

```

BASE COUNT 44147 a 43012 c 43431 g 43839 t 3315 others
ORIGIN

```

```

Query Match 54.9% Score 417, DB 2, Length 177744;
Best Local Similarity 91.0% Pred. No. 6,6e-55;

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model1

Run on: March 28, 2003, 17:25:48 / Search time 2063.01 Seconds

(without alignments)
8972.025 Million cell updates/sec

Title: US-09-867-274-3

Perfect score: 636
Sequence: 1 atgcagccctcactagcccc.....agctgagagcctcactag 636

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenBankl: *
1: gb_ba: *
2: gb_hcg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pac: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_scs: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_scs: *
28: em_un: *
29: em_vl: *
30: em_hcg_hum: *
31: em_hcg_inv: *
32: em_hcg_other: *
33: em_hcg_mus: *
34: em_hcg_pln: *
35: em_hcg_rod: *
36: em_hcg_mam: *
37: em_hcg_vtc: *
38: em_sy: *
39: em_hcgo_hum: *
40: em_hcgo_mus: *
41: em_hcgo_other: *

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	636	100.0	636	6	AX323455 Sequence
2	634.4	99.7	638	10	AF326740 Mus muscu
3	557.6	87.7	674	10	AF326741 Rattus no
4	504.8	79.4	642	9	AF326742 Cercopit
5	501.6	78.9	759	6	AX323453 Sequence
6	501.6	78.9	2271	6	AX342535 Sequence
7	501.6	78.9	2296	9	AF331844 Homo sapi
8	501.6	78.9	2323	9	AF326739 Homo sapi
9	501.6	78.9	2329	6	AX056687 Sequence
10	426.6	67.1	81806	10	AF326737 Mus muscu
11	426.6	67.1	110000	2	AC068782-2 Continuation (3 of
12	426.6	67.1	205277	2	AC012296 Mus muscu
13	426.6	67.1	208135	10	AC068807 Mus muscu
14	425	66.8	198508	10	AL591145 Mus muscu
15	415.8	65.4	532	4	AF326738 Bos tauru
16	409	64.3	101804	2	AC098160 Rattus no
17	409	64.3	104898	2	AC121721 Rattus no
18	367	57.7	7099	6	AX342537 Sequence
19	367	57.7	21501	9	AF326736 Homo sapi
20	367	57.7	80117	9	AC055813 Homo sapi
21	367	57.7	94752	9	AC003098 Homo sapi
22	282.6	44.4	177744	2	AC073954 Homo sapi
23	214.4	33.7	51575	2	AC023810 Mus muscu
24	81.2	13.1	968	10	BC021458 Mus muscu
25	81.2	12.8	220967	2	AC124776 Mus muscu
26	71.6	11.3	182253	2	AC117353 Rattus no
27	67.2	10.6	900	6	AF017832 Sequence
28	67.2	10.6	1091	6	AX454484 Sequence
29	67.2	10.6	1091	6	AX490962 Sequence
30	67.2	10.6	1723	9	BC008484 Homo sapi
31	65.6	10.3	618	6	AX429979 Sequence
32	65.6	10.3	621	6	AX429978 Sequence
33	65.6	10.3	1692	6	AX429953 Sequence
34	65.2	10.3	2409	9	AK093408 Homo sapi
35	65.2	10.3	125372	9	AC079155 Homo sapi
36	63.6	10.0	1500	9	AL050024 Homo sapi
37	59.2	9.3	32748	1	AB070951 Streptomy
38	59.2	9.3	37638	2	AC111799 Rattus no
39	59	9.3	78862	2	AC095361 Rattus no
40	57.4	9.0	172490	2	AC117097 Rattus no
41	57	9.0	412	6	AX071502 Sequence
42	55.2	8.7	125020	9	AF429315 Homo sapi
43	54.6	8.6	92055	2	AC105692 Rattus no
44	54.4	8.6	132884	2	AC128403 Rattus no
45	54.4	8.6	214885	2	AC094318 Rattus no

ALIGNMENTS

RESULT 1
AX323455
LOCUS AX323455
DEFINITION Sequence 3 from Patent WO01923308.
ACCESSION AX323455
VERSION AX323455.1 GI:18094217
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Paszty, C. J. and Gao, Y.
TITLE Cytidine-knot polypeptides: cloaked-2 molecules and uses thereof
JOURNAL Patent: WO 01923308-A 3 06-DEC-2001;

Pred. No. is the number of results predicted by chance to have a

Amgen, Inc. (US)
 FEATURES Location/Qualifiers
 Source 1..636
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 BASE COUNT 114 a 224 c 207 g 91 t
 ORIGIN

Query Match 100.0%; Score 636; DB 6; Length 636;
 Best Local Similarity 100.0%; Pred. No. 4.8e-99;
 Matches 636; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGAGCCCTCACTAGCCGCTGCTCATCTGCTTGTGACAGCTGCTTGTGCT 60
 1 ATGAGCCCTCACTAGCCGCTGCTCATCTGCTTGTGACAGCTGCTTGTGCT 60
 61 GTGAGGGGCGAGGGGTGGCAAGCTTTCAGAAATGATCCACAGAGTCAATCCAGGGCTT 120
 61 GTGAGGGGCGAGGGGTGGCAAGCTTTCAGAAATGATCCACAGAGTCAATCCAGGGCTT 120
 121 GGAAGATACCCCGAGGCTCTCTGAGAACACCAACATGAAACCGGGGGAATGGA 180
 121 GGAAGATACCCCGAGGCTCTCTGAGAACACCAACATGAAACCGGGGGAATGGA 180
 181 GCGAGACCTCCCGACCACTCCCTATGAGCCCAAGATGTGTCCAGTACAGTCCGCGAG 240
 181 GCGAGACCTCCCGACCACTCCCTATGAGCCCAAGATGTGTCCAGTACAGTCCGCGAG 240
 241 CTGCACTACACCCGCTTCTTGACAGACGCGCCATGCGGAGCGCAAGCGGTCACCGAG 300
 241 CTGCACTACACCCGCTTCTTGACAGACGCGCCATGCGGAGCGCAAGCGGTCACCGAG 300
 301 TTGGTGTGCTCCCGGCGAGTGGGCGCGCGCGGCTGCTGCGCAACCGCATCGGGCGGTG 360
 301 TTGGTGTGCTCCCGGCGAGTGGGCGCGCGCGGCTGCTGCGCAACCGCATCGGGCGGTG 360
 361 AAGTGTGGGCGCGGAGCGGACCGGATTTCCGTGATCCCGATCGCTACCGCGCGAG 420
 361 AAGTGTGGGCGCGGAGCGGACCGGATTTCCGTGATCCCGATCGCTACCGCGCGAG 420
 421 CGGGTGCAGCTGCTGTGCGCGCGGGGCGCGCGCGGCTGCGCAAGGTGCGTCTGGTG 480
 421 CGGGTGCAGCTGCTGTGCGCGCGGGGCGCGCGCGGCTGCGCAAGGTGCGTCTGGTG 480
 481 GCTCTGTGCAAGTCAAGCGGCTTCCACCAACCAAGTGTGAGGACTTTC 540
 481 GCTCTGTGCAAGTCAAGCGGCTTCCACCAACCAAGTGTGAGGACTTTC 540
 541 GGGCGGAGACCGGCGCGGCGGAGAGGAGTGTGCAAGCGCGGCGCGCGCGGAGGCC 600
 541 GGGCGGAGACCGGCGCGGCGGAGAGGAGTGTGCAAGCGCGGCGCGCGCGGAGGCC 600
 601 AAAGCCAAACGAGCGGAGCTGAGAACGCTTACTAG 636
 601 AAAGCCAAACGAGCGGAGCTGAGAACGCTTACTAG 636
 RESULT 2
 AF326740 638 bp mRNA linear ROD 28-FEB-2001
 LOCUS Mus musculus sclerostin mRNA, complete cds.
 DEFINITION AF326740
 ACCESSION AF326740.1 GI:13161022
 VERSION
 KEYWORDS
 SOURCE Mus musculus.
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 638)
 Brunrow, M.E., Gardner, J.C., Van Ness, J., Paepers, B.W.,
 Kovacevich, B.R., Proll, S., Skonier, J.E., Zhao, L., Sabo, P.J.,
 Fu, Y.H., Altsch, R.S., Gillet, L., Colbert, T., Tacconi, P., Galas, D.,
 Hamerema, H., Beighton, P. and Mulligan, J.T.

TITLE Bone dysplasia sclerosteosis results from loss of the SOST gene product, a novel cystine knot-containing protein
 JOURNAL Am. J. Hum. Genet. 68 (3), 577-589 (2001)
 MEDLINE 21090529
 PUBMED 11179006
 REFERENCE 2 (bases 1 to 638)
 AUTHORS Brunrow, M.E., Gardner, J.C., Van Ness, J., Paepers, B.W., Kovacevich, B.R., Proll, S., Skonier, J.E., Zhao, L., Sabo, P.J., Fu, Y.H., Altsch, R.S., Gillet, L., Colbert, T., Tacconi, P., Galas, D., Hamerema, H., Beighton, P. and Mulligan, J.T.
 Direct Submission
 JOURNAL Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631 220th St. SE, Bothell, WA 98021, USA

FEATURES Location/Qualifiers
 Source 1..638
 /organism="Mus musculus"
 /strain="129Sv/J"
 /db_xref="taxon:10090"
 /chromosome="11"
 1..636
 /codon_start=1
 /product="sclerostin"
 /protein_id="AAK13455.1"
 /db_xref="GI:13161023"
 /translation="MQPSLAPCLICLVHNAFCVAVGGQWQAFRNDATVETPGLGEYP
 EPPENNQTMRAENGSRPHHPYDAKGVSEYSCRELYHTRFLTDGPRSAKVTLY
 CGGCGPRLPLPAIGRWKWRPNRPFRCI PRYBAROROLLCPGAAPRSKRVLV
 ASCKRULTFPHNSELKDPRPTARQKRRKRPARGAKNAQALENAV"

CDS

BASE COUNT 114 a 224 c 209 g 91 t
 ORIGIN
 Query Match 99.7%; Score 634.4; DB 10; Length 638;
 Best Local Similarity 99.8%; Pred. No. 9e-99; 1; Indels 0; Gaps 0;
 Matches 635; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 1 ATGAGCCCTCACTAGCCGCTGCTCATCTGCTTGTGACAGCTGCTTGTGCT 60
 1 ATGAGCCCTCACTAGCCGCTGCTCATCTGCTTGTGACAGCTGCTTGTGCT 60
 61 GTGAGGGGCGAGGGGTGGCAAGCTTTCAGAAATGATCCACAGAGTCAATCCAGGGCTT 120
 61 GTGAGGGGCGAGGGGTGGCAAGCTTTCAGAAATGATCCACAGAGTCAATCCAGGGCTT 120
 121 GGAAGATACCCCGAGGCTCTCTGAGAACACCAACATGAAACCGGGGGAATGGA 180
 121 GGAAGATACCCCGAGGCTCTCTGAGAACACCAACATGAAACCGGGGGAATGGA 180
 181 GCGAGACCTCCCGACCACTCCCTATGAGCCCAAGATGTGTCCAGTACAGTCCGCGAG 240
 181 GCGAGACCTCCCGACCACTCCCTATGAGCCCAAGATGTGTCCAGTACAGTCCGCGAG 240
 241 CTGCACTACACCCGCTTCTTGACAGACGCGCCATGCGGAGCGCAAGCGGTCACCGAG 300
 241 CTGCACTACACCCGCTTCTTGACAGACGCGCCATGCGGAGCGCAAGCGGTCACCGAG 300
 301 TTGGTGTGCTCCCGGCGAGTGGGCGCGCGCGGCTGCTGCGCAACCGCATCGGGCGGTG 360
 301 TTGGTGTGCTCCCGGCGAGTGGGCGCGCGCGGCTGCTGCGCAACCGCATCGGGCGGTG 360
 361 AAGTGTGGGCGCGGAGCGGACCGGATTTCCGTGATCCCGATCGCTACCGCGCGAG 420
 361 AAGTGTGGGCGCGGAGCGGACCGGATTTCCGTGATCCCGATCGCTACCGCGCGAG 420
 421 CGGGTGCAGCTGCTGTGCGCGGGGCGCGCGCGGCTGCGCAAGGTGCGTCTGGTG 480
 421 CGGGTGCAGCTGCTGTGCGCGGGGCGCGCGCGGCTGCGCAAGGTGCGTCTGGTG 480
 481 GCTCTGTGCAAGTCAAGCGGCTTCCACCAACCAAGTGTGAGGACTTTC 540
 481 GCTCTGTGCAAGTCAAGCGGCTTCCACCAACCAAGTGTGAGGACTTTC 540
 541 GGGCGGAGACCGGCGCGGCGGAGAGGAGTGTGCAAGCGCGGCGCGCGGAGGCC 600
 541 GGGCGGAGACCGGCGCGGCGGAGAGGAGTGTGCAAGCGCGGCGCGCGGAGGCC 600

QY	601	AAAGCAACACGCGGAGCTGAGAAAGCCCTACTAG	636
Db	601	AAAGCAACACGCGGAGCTGAGAAAGCCCTACTAG	636
RESULT 3			
LOCUS	AF326741	674 bp	mRNA linear
DEFINITION	Rattus norvegicus sclerostin mRNA, complete cds.		ROD 28-FEB-2001
ACCESSION	AF326741		
VERSION	AF326741.1	GI:13161025	
KEYWORDS			
SOURCE			
ORGANISM	Rattus norvegicus.		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
REFERENCE	1 (bases 1 to 674)		
AUTHORS	Brunkow, M.E., Gardner, J.C., Van Ness, J., Paeppe, B.W., Kovacevich, B.R., Prohl, S., Skonier, J.E., Zhao, L., Sabo, P.J., Fu, Y.H., Altsch, R.S., Gillett, L., Colbert, T., Tacconi, P., Galas, D., Hamersma, H., Beighton, P. and Mulligan, J.T.		
TITLE	Bone dysplasia sclerosteosis results from loss of the SOST gene product, a novel cysteine knot-containing protein		
JOURNAL	Am. J. Hum. Genet. 68 (3), 577-589 (2001)		
MEDLINE	21090529		
PUBMED	11179006		
REFERENCE	2 (bases 1 to 674)		
AUTHORS	Brunkow, M.E., Gardner, J.C., Van Ness, J., Paeppe, B.W., Kovacevich, B.R., Prohl, S., Skonier, J.E., Zhao, L., Sabo, P.J., Fu, Y.H., Altsch, R.S., Gillett, L., Colbert, T., Tacconi, P., Galas, D., Hamersma, H., Beighton, P. and Mulligan, J.T.		
TITLE	Direct Submission		
JOURNAL	Submitted (07-DEC-2000) Genomics, Celtech Neuroscience Inc., 1631		
FEATURES	220th St. SE, Bothell, WA 98021, USA		
SOURCE	location/Qualifiers		
	1..674		
	/organism="Rattus norvegicus"		
	/strain="Sprague-Dawley"		
	/db_xref="taxon:10116"		
	1..32		
	33..674		
	/codon_start=1		
	/product="sclerostin"		
	/protein_id="AAK13456.1"		
	/db_xref="GI:13161026"		
	/translation="MQLSLAPACLCLVHAFAVAVESKQMAFNKATETIIPLGLREYPLPGEENNOTMNAENGRRPHRPHYDVKVSESCGLHYTRPMGPRSAKPYLELVSGGCGPARLLPNAIGRYKWRPNPDRPCIDRIFAQRVGLTCGGAAPSRRKRLVASCCKRRLRFRNOSBLDKFGETIRAPQGRKRPFAAGAKANQALENAY"		
BASE COUNT	128 a 240 c 208 g 98 t		
ORIGIN			
Query Match	87.7%, Score 557.6; DB 10; Length 674;		
Best Local Similarity	93.0%; Pred. No. 1..1e-85;		
Matches	597, Conservative 0; Mismatches 39; Indels 6; Gaps 1;		
QY	1	ATGCAGCCTCACTAGCCCCGGTCTCATCTGGCTACTTGTGACGCTGCTTGTCCT	60
Db	33	ATGCAGCCTCACTAGCCCCGGTCTCTGCTTCCCTGCTTGTATGACGCTTGTCCT	92
QY	61	GTGAGAGCCGAGGGGTGGCAAGCTTTCAGAAATGTCACAGAGGTGATCCCGAGGCTT	120
Db	93	GTGAGAGCCGAGGGGTGGCAAGCTTCAAGATATGTCACAGAAATATCTCCGGAATC	152
QY	121	GGAGAGTACCCCGAGCTCTT-----CTTAAAGCAACCAAGCCATGTAACCGGCGAG	174
Db	153	AGAGAGTACCAAGAGCTCTCTCAGGAATTAAGAAACAACCAATGTAACCGGCGAG	212
QY	175	AATGAGGAGAGACCTCCCAACATCCCTATATAGCGCAAGATGTGTCCGACTTACAGCTGC	234

Db	213	AAAGGAGGAGAGACCCCCCAACCATCTTATGACACCAAGAGTGTCCGAGTACAGCTGC	272
QY	235	GGGAGGCTGCACTAACACCCGCTTCTTGACAGACGGCCCATGCGGACGACCGCCGTC	294
Db	273	CGGAGGCTGCACTAACACCCGCTTCTTGACAGACGGCCCGTTCGACGTCACGACCGGTC	332
QY	295	ACCGAGTTGTGTGCTCCGAGCCAGTGCAGCCCGCGCGAGCTGCTGCCACGACATCCGG	354
Db	333	ACCGAGTTGTGTGCTCCGAGCCAGTGCAGCCCGCGCGAGCTGCTGCCACGACATCCGG	392
QY	355	CGCGTGAAGTGTGGCGCCCGAACGACCGAATTTCCGCTGCATCCCGATCGTACCGC	414
Db	393	CGCGTGAAGTGTGGCGCCCGAACGACCGAATTTCCGCTGCATCCCGATCGTACCGC	452
QY	415	GGCGAGCGGAGTGCAGCTGTGACCCCGGAGGCGGAGGCGCGCGTCCCGGACAGGTGGT	474
Db	453	GGCGAGCGGAGTGCAGCTGTGACCCCGGAGGCGGAGGCGCGCGTCCCGGACAGGTGGT	512
QY	475	CTGGTGGCTCTGTGCAAGTGCAGAGCGCTCAACCCGCTTCCACACAGTGCAGGCTCAAG	534
Db	513	CTGGTGGCTCTGTGCAAGTGCAGAGCGCTCAACCCGCTTCCACACAGTGCAGGCTCAAG	572
QY	535	GACCTCGGAGCGGAGACCGCGCGCGCCGACAGAGGCTCCGACAGCCGCGCGCGCGG	594
Db	573	GACCTCGGAGCGGAGACCGCGCGCGCCGACAGAGGCTCCGACAGCCGCGCGCGCGG	632
QY	595	GGAGCCAAAGCCAAACAGAGCGGAGCTGAGAGACGCTACTAG	636
Db	633	GGAGCCAAAGCCAAACAGAGCGGAGCTGAGAGACGCTACTAG	674
RESULT 4	AF326742	642 bp	linear PRI 28-FEB-2001
LOCUS	AF326742	Cercopithecus aethiops sclerostin mRNA, complete cds.	
DEFINITION	AF326742		
ACCESSION	AF326742.1	GI:13161028	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE		</	

QY 595 GGAGCCAAAGCCAGGCGGAGCTGGAGAAAGCCTACTAG 636
 Db 638 AGGCCAAAGCCAGGCGGAGCTGGAGAAAGCCTACTAG 679
 RESULT 8
 LOCUS AF326739 2323 bp mRNA linear PRI 28-FEB-2001
 DEFINITION Homo sapiens sclerostin mRNA, complete cds.
 ACCESSION AF326739
 VERSION AF326739.1 GI:13161019
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 2323)
 AUTHORS Brunkow, M.E., Gardner, J.C., Van Ness, J., Paepker, B.W.,
 Kovacevich, B.R., Prohl, S., Skonier, J.B., Zhao, L., Sabo, P.J.,
 Fu, Y.H., Altsch, R.S., Gillett, L., Colbert, T., Tacconi, P., Galas, D.,
 Hamersma, H., Beighton, P. and Mulligan, J.T.
 Bone dysplasia sclerosteosis results from loss of the SOST gene
 product, a novel cysteine knot-containing protein
 Am. J. Hum. Genet. 68 (3), 577-589 (2001)
 JOURNAL MEDLINE 21090529
 PUBLISHED 11/19/06
 REFERENCE 2 (bases 1 to 2323)
 AUTHORS Brunkow, M.E., Gardner, J.C., Van Ness, J., Paepker, B.W.,
 Kovacevich, B.R., Prohl, S., Skonier, J.B., Zhao, L., Sabo, P.J.,
 Fu, Y.H., Altsch, R.S., Gillett, L., Colbert, T., Tacconi, P., Galas, D.,
 Hamersma, H., Beighton, P. and Mulligan, J.T.
 Direct Submission
 Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631
 220th St. SE, Bothell, WA 98021, USA
 JOURNAL 220th St. SE, Bothell, WA 98021, USA
 FEATURES
 source
 1. .2323
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="17"
 /map="17q21"
 1. .47
 48. .689
 /codon_start=1
 /product="sclerostin"
 /protein_id="AAK13454.1"
 /translation="MQLPLALCLVCLLVHTARVVGQGMQAFKNDATETIIPELGEYP
 EPPLENNKTNRAENGSRPHHPETDVSSEYSCREHFTRYVTDGPRRAKPVTE
 LVSGGCGPARLFPNAIGRKMMRSPGPRCPIDRYRAORVOLLCPGGEAPRAKVR
 LVASCKCKILTRRHNSLSEKDPTEAARQKRRPRARSAKANQAELENNY"
 690. .2323
 BASE COUNT 590 a 635 c 614 g 484 t
 ORIGIN
 Query Match 78.9%; Score 501.6; DB 9; Length 2323;
 Best Local Similarity 87.5%; Pred. No. 3.1e-76;
 Matches 562; Conservative 0; Mismatches 74; Indels 6; Gaps 1;

Db 228 AACGAGGCGGCGCTCCCAACCAACCCCTTGAAGACAAAGAGTGTCCGATACAGCTGC 287
 QY 225 CGCAGCTGCACTACACCCGCTTCTGACAGACGCGCCATAGCGGCAAGCGGCTC 294
 Db 288 CGCAGCTGCACTTACACCCGCTTCTGACAGACGCGCCATAGCGGCAAGCGGCTC 347
 QY 295 ACCGAGTTGTGTGCTCCGAGCAAGTGGGCGCGCGGCTGCTGCCAAGCATCGAGG 354
 Db 348 ACCGAGTTGTGTGCTCCGAGCAAGTGGGCGCGCGGCTGCTGCCAAGCATCGAGG 407
 QY 355 CGCAGTTGTGTGCTCCGAGCAAGTGGGCGCGCGGCTGCTGCCAAGCATCGAGG 414
 Db 408 CGCAGTTGTGTGCTCCGAGCAAGTGGGCGCGCGGCTGCTGCCAAGCATCGAGG 467
 QY 415 GCGGCGGCGGCTGCAAGTGGGCGCGCGGCTGCTGCCAAGCATCGAGGCTGCT 474
 Db 468 GCGGCGGCGGCTGCAAGTGGGCGCGCGGCTGCTGCCAAGCATCGAGGCTGCT 527
 QY 475 CTGCTGCGCTGCTGCAAGTGGGCGCGCGGCTGCTGCCAAGCATCGAGGCTGCT 534
 Db 528 CTGCTGCGCTGCTGCAAGTGGGCGCGCGGCTGCTGCCAAGCATCGAGGCTGCT 587
 QY 535 GACTTGGGCGGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 594
 Db 588 GACTTGGGCGGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 647
 QY 595 GGAGCCAAAGCCAGGCGGAGCTGGAGAAAGCCTACTAG 636
 Db 648 AGGCCAAAGCCAGGCGGAGCTGGAGAAAGCCTACTAG 689
 RESULT 9
 LOCUS AX056687 2329 bp DNA linear PAT 17-JUN-2001
 DEFINITION Sequence 19 from Patent WO0075317.
 ACCESSION AX056687
 VERSION AX056687.1 GI:12309667
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 2329)
 AUTHORS Wood, W.I., Goddard, A., Gurney, A.L., Smith, V., Watanabe, C.K. and
 Wood, W.I.
 Compositions and methods for the treatment of tumor
 Patent: WO 0075317-A 19 14-DEC-2000;
 JOURNAL Genentech, Inc. (US)
 FEATURES
 source
 1. .2329
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 587 a 645 c 612 g 485 t
 ORIGIN
 Query Match 78.9%; Score 501.6; DB 6; Length 2329;
 Best Local Similarity 87.5%; Pred. No. 3.1e-76;
 Matches 562; Conservative 0; Mismatches 74; Indels 6; Gaps 1;

Db 242 AACGAGGCGGCGCTCCACCACCCCTTTGAGACAAAGAGCTGTCCAGTACAGCTGC 301

Qy 235 CGCGAGCTGCACTAGACCCGCTTCTCTGACAGACGCGCCATGCGGACG3CCAAAGCGCTC 234

Db 302 CGCGAGCTGCACTAGACCCGCTTCTCTGACAGACGCGCCATGCGGACG3CCAAAGCGCTC 301

Qy 235 ACCGAGTTGTGTGCTCCGCGGACGAGTGCAGCCCGCGGCGCTGCTGCCAAAGCCATCGAG 354

Db 362 ACCGAGCTGTGTGTCTCCGCGGACGAGTGCAGCCCGCGGCGCTGCTGCCAAAGCCATCGAG 421

Qy 355 CGCGGAGAGTGTGTGCTCCGCGGACGAGATTTTCCGCTGATCCCGGATCCGCTACCGC 414

Db 422 CGCGGAGAGTGTGTGCTCCGCGGACGAGATTTTCCGCTGATCCCGGATCCGCTACCGC 481

Qy 415 GCGCAGCGGAGTGCAGCTGTGCTCCGCGGACGAGATTTTCCGCTGATCCCGGATCCGCTACCGC 474

Db 482 GCGCAGCGGAGTGCAGCTGTGCTCCGCGGACGAGATTTTCCGCTGATCCCGGATCCGCTACCGC 511

Qy 475 CTGTGTGCTGTGTGCTGCAAGTGCAGCGCTTCAACCGATCGAGCTCAAG 534

Db 542 CTGTGTGCTGTGTGCTGCAAGTGCAGCGCTTCAACCGATCGAGCTCAAG 601

Qy 535 GACTTTCGCGCGGAGACCGCGGCGCGGAGAGGCTGCGAAAGCGCGGCGCGCGCGCGCG 594

Db 602 GACTTTCGCGCGGAGACCGCGGCGCGGAGAGGCTGCGAAAGCGCGGCGCGCGCGCGCG 661

Qy 595 GGAGCCAAAGCCAAACGAGCGGAGCTGAGAGAGCGCTACTAG 636

Db 662 AGCGCCAAAGCCAAACGAGCGGAGCTGAGAGAGCGCTACTAG 703

RESULT 10

AF326737 81806 bp DNA linear ROD 15-APR-2002

LOCUS AF326737

DEFINITION Mus musculus sclerostin (Sost) gene, complete cds; and Meox1 gene, partial sequence.

ACCESSION AF326737

VERSION AF326737.2 GI:20150447

KEYWORDS

SOURCE Mus musculus.

ORGANISM Mus musculus; Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

AUTHORS Brunkow, M.E., Gardner, J.C., Van Ness, J., Paepker, B.W., Koyavevich, B.R., Prohl, S., Skonier, J.E., Zhao, L., Sabo, P.J., Fu, Y.H., Alish, R.S., Gillet, L., Colbert, T., Tacconi, P., Galas, D., Hamersma, H., Beighton, P. and Mulligan, J.T.

TITLE Bone dysplasia sclerosteosis results from loss of the SOST gene product, a novel cysteine knot-containing protein

JOURNAL Am. J. Hum. Genet. 68 (3), 577-589 (2001)

MEDLINE 21090529

PUBMED 11179006

REFERENCE

AUTHORS Staehling-Hampton, K., Prohl, S., Paepker, B.W., Zhao, L., Charnley, P., Brown, A., Gardner, J.C., Galas, D., Schatzman, R.C., Beighton, P., Papadopoulos, S., Hamersma, H. and Brunkow, M.E.

TITLE A 52 kb deletion in the SOST - MEOX1 intergenic region on 17q12-q21 is associated with van Buchem disease in the Dutch population

JOURNAL Unpublished

REFERENCE

AUTHORS Brunkow, M.E., Gardner, J.C., Van Ness, J., Paepker, B.W., Koyavevich, B.R., Prohl, S., Skonier, J.E., Zhao, L., Sabo, P.J., Fu, Y.H., Alish, R.S., Gillet, L., Colbert, T., Tacconi, P., Galas, D., Hamersma, H., Beighton, P. and Mulligan, J.T.

TITLE Direct Submission

JOURNAL Submitted (07-DEC-2000) Genomics, Celltech Chiroscience Inc., 1631 220th St. SE, Bothell, WA 98021, USA

REFERENCE

AUTHORS Staehling-Hampton, K., Prohl, S., Paepker, B.W., Zhao, L., Charnley, P., Brown, A., Gardner, J.C., Galas, D., Schatzman, R.C., Beighton, P., Papadopoulos, S., Hamersma, H. and Brunkow, M.E.

TITLE Direct Submission

JOURNAL Submitted (02-AUG-2001) Genomics, Celltech Red Inc., 1631 220th St SE, Bothell, WA 98021, USA

REMARK Sequence update by submitter

COMMENT On Apr 15, 2002 this sequence version replaced gi:13161013.

FEATURES

source

1. 81806

/organism="Mus musculus"

/strain="129SV/7"

/db_xref="taxon:10090"

/chromosome="11"

/map="between Sost and Meox1"

<8036..>11163

/gene="Sost"

join(<8036..8249,10742..>11163)

/gene="Sost"

/product="sclerostin"

join(8036..8249,10742..11163)

/gene="Sost"

/codon_start=1

/product="sclerostin"

/protein_id="AAK13452.1"

/db_xref="GI:13161014"

/translation="MQSLAPCLICLVHAFCAVEGGQWAFRDATETVIGLGEYEP

EPPENNQTMNRAENGRPPHPYDAKDVSEYSEELHYTRFLTDPGRSAKPTVELV

SGCCGPARLPVIAIGRVKMRPNGDFPCIDRYARAVOLLCPGAAPSRKRVRLV

ASCKCKRLTRFPHNOSLNDPQEPETARPQGRKPRFGAARAKNOMLENAI"

81492..>81806

/gene="Meox1"

/db_xref="taxon:10090"

/gene="Meox1"

/product="Meox1"

BASE COUNT 20982 a 20323 c 20010 g 20491 t

ORIGIN

Query Match 67.1%; Score 426.6; DB 10; Length 81806;

Best Local Similarity 95.8%; Pred. No. 1e-63;

Matches 438; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 180 AGGAGACCTCCCGACCATTCCTATGACGCGAAAGATGTCGAGTACAGTGCAGCA 239

Db 10707 AGGAGACCTCCCGACCATTCCTATGACGCGAAAGATGTCGAGTACAGTGCAGCA 10766

Qy 240 GTCGACTACACCCGCTTCTCTGACAGACGCGCCATGCGGACGCGCCAGCGCTACCGCA 299

Db 10767 GTCGACTACACCCGCTTCTCTGACAGACGCGCCATGCGGACGCGCCAGCGCTACCGCA 10826

Qy 300 GTTGTGTGTCTCCGCGGACGAGTGCAGCCCGCGGCGCTGTGCTGCCAAGCCATCGGCGCT 359

Db 10827 GTTGTGTGTCTCCGCGGACGAGTGCAGCCCGCGGCGCTGTGCTGCCAAGCCATCGGCGCT 10886

Qy 360 GAAAGTGTGCGCGCGCGGACGAGATTTCCGCTGACATCCGCGATCGCTACCGCGCGCA 419

Db 10887 GAAAGTGTGCGCGCGCGGACGAGATTTCCGCTGACATCCGCGATCGCTACCGCGCGCA 10946

Qy 420 GCGGATGACAGCTGTGCTCCGCGGCGCGCGCGCGCGCGCTGCGCGCAAGTGCCTGTGT 479

Db 10947 GCGGATGACAGCTGTGCTCCGCGGCGCGCGCGCGCGCGCTGCGCGCAAGTGCCTGTGT 11006

Qy 480 GCGCTGTGTAATGTGCAAGCGCTTCAACCGATCGAGTGCAGTGCAGTGCAGTGCAGT 539

Db 11007 GCGCTGTGTAATGTGCAAGCGCTTCAACCGATCGAGTGCAGTGCAGTGCAGTGCAGT 11066

Qy 540 CGGCGCGGAGACCGCGCGCGCGCGGAGAGGAGTGCAGAGCGCGCGCGCGCGCGCGCGCGCG 599

Db 11067 CGGCGCGGAGACCGCGCGCGCGCGGAGAGGAGTGCAGAGCGCGCGCGCGCGCGCGCGCGCG 11126

Qy 600 CAAAGCCAAACGAGCGGAGCTGAGAGAGCGCTTACTAG 636

Db 11127 CAAAGCCAAACGAGCGGAGCTGAGAGAGCGCTTACTAG 11163

RESULT 11

AC068782_2

WPCOMMENT

Sequence split into 5 fragments LOCUS AC068782 Accession AC068782

Fragment Name	Begin	End
AC068782_0	1	110000
AC068782_1	100001	210000
AC068782_2	200001	310000
AC068782_3	300001	410000
AC068782_4	400001	411130

Continuation (3 of 5) of AC068782 from base 200001 (AC068782 Mus musculus chromosome 11)

Query Match 67.1%; Score 426.6; DB 2; Length 110000;

Best Local Similarity 95.8%; Pred. No. 9.7e-64;

Matches 438; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

DB	Accession	Score	DB	Length
QY	180 AGGACAGCTCTCCACCATCTCTATGACCGCAAGATGTGTCCGAGTACAGCTGCGCGCA	239		
DB	56567 AGGCTGACAGCCCTTCAAGCATCTCTCTCCGAGATGTGTCCGAGTACAGCTGCGCGCA	56626		
QY	240 GGTGCACTACACCCGCTTCTGACAGAGCCCATGCGGAGCCGAGCCGATCACCGA	299		
DB	56627 GGTGCACTACACCCGCTTCTGACAGAGCCCATGCGGAGCCGAGCCGATCACCGA	56686		
QY	300 GTTGTGTGCTCCGCGCAAGTCCGCGCGCGCGCTGTCTGCCAAGCCATCCGCGCGCT	359		
DB	56687 GTTGTGTGCTCCGCGCAAGTCCGCGCGCGCGCTGTCTGCCAAGCCATCCGCGCGCT	56746		
QY	360 GAAGTGTGCGCGCGCGCAAGCCGATTTCCGCTGCAATCCGAGTACCGCGCGCA	419		
DB	56747 GAAGTGTGCGCGCGCGCAAGCCGATTTCCGCTGCAATCCGAGTACCGCGCGCA	56806		
QY	420 GCGGAGTGTGCTGTGTGCGCGCGCGCGCGCGCGCTGTCTGCCAAGTGTGTGT	479		
DB	56807 GCGGAGTGTGCTGTGTGCGCGCGCGCGCGCGCGCTGTCTGCCAAGTGTGTGT	56866		
QY	480 GGCCTGTGCAATGTCAAGCGCTCTACCCGCTTCCACATCACTCGGAGCTCAAGACTT	539		
DB	56867 GGCCTGTGCAATGTCAAGCGCTCTACCCGCTTCCACATCACTCGGAGCTCAAGACTT	56926		
QY	540 CGGCGCGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGTCTGCCAAGTGTGT	599		
DB	56927 CGGCGCGCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGTCTGCCAAGTGTGT	56986		
QY	600 CAAAGCCAAACGAGCGAGCTGTGAGAAAGCGCTACTAG	636		
DB	56987 CAAAGCCAAACGAGCGAGCTGTGAGAAAGCGCTACTAG	57023		

RESULT 12
AC012296
LOCUS
DEFINITION Mus musculus chromosome 11 clone RP23-346P7 map 11, *** SEQUENCING
IN PROGRESS ***; 10 unordered pieces.
AC012296 205277 bp DNA linear HTG 16-AUG-2002
AC012296 10 GI:22267737
VERSION HTG; HTGS_PBASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 205277)
Birren, B., Nusbaum, C. and Lander, E.
Mammalian genome project: A complete sequence of the mouse genome.
Unpublished
2 (bases 1 to 205277)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Baldwin, J., Barna, N., Beckwith, R., Boguslavsky, L., Bouhassira, B., Brown, A., Castle, A., Colangelo, M., Collins, S., Collins, A., Cooke, P., Dearfield, K., Dewar, K., Dominko, M., Donnelly, L., Doyle, M., Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D., Galagan, J., Galagan, S., Grant, G., Hages, B., Heath, A., Horton, L., Howland, J. C., Johnson, R., Jones, C., Kamm, L., Karas, A., Klein, J., Lehotzky, J., Lien, C., Locke, K., Macdonald, P., Marquis, N., McSwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J.,

TITLE
JOURNALREFERENCE
AUTHORSTITLE
JOURNAL
COMMENT

Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Teste, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X., Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

Submitted (22-OCT-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 205277)

Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Bouhassira, B., Camarata, J., Chang, J., Chazaro, B., Choepel, T., Collins, A., Cooke, A., Cooke, P., Dearfield, K., Dewar, K., Diaz, U. S., Dodge, S., Fero, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Galagan, S., Gord, S., Graham, L., Grand-Pierre, N., Hages, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrum, J., Menes, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunhah, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Teste, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission

Submitted (16-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Aug 16, 2002 this sequence version replaced gi:20177756.

All repeats were identified using RepeatMasker:

Smith, A. F. A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: MIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L3205

Center clone name: 346_P_7

* NOTE: This is a 'working draft' sequence. It currently consists of 10 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 92178: contig of 92178 bp in length

* 92179 92278: gap of 100 bp

* 92279 128704: contig of 36426 bp in length

* 128705 128804: gap of 100 bp

* 128805 140272: contig of 11468 bp in length

* 140273 140372: gap of 100 bp

* 140373 147001: contig of 6629 bp in length

* 147002 147101: gap of 100 bp

* 147102 170176: contig of 23075 bp in length

* 170177 170276: gap of 100 bp

* 170277 173609: contig of 3333 bp in length

* 173610 173709: gap of 100 bp

* 173710 175722: contig of 2013 bp in length

* 175723 175822: gap of 100 bp

* 175823 181949: contig of 6127 bp in length

* 181950 182049: gap of 100 bp

* 182050 191157: contig of 9108 bp in length

* 191158 191257: gap of 100 bp

* 191258 205277: contig of 14020 bp in length.

Location/Qualifiers

1..205277

/organism="Mus musculus"


```

/db_xref="taxon:10090"
/chromosome="11"
/map="11"
/clone_1ib="RP23-346P7"
BASE COUNT 53425 a 49464 c 50390 g 50980 t 1018 others
ORIGIN

```

```

Query Match 67.1%; Score 426.6; DB 2; Length 205277;
Best Local Similarity 95.8%; Pred. No. 8,7e-64;
Matches 438; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

```

```

QY 180 AGGCAAGCTCTCCACCATCTATGACGCAAGATGTGTCCGATGACAGTCCGCGA 239
DB 57833 AGGCTGAGCCCTTCAAGCATCTCTCTCCGAGATGTGTCCGAGTACAGTCCGCGA 57892
QY 240 GCGCACTACACCCGCTTCTTCAAGACGCGCCATGCGCGCAAGCGCGTCAACCA 299
DB 57893 GCTGCACTACACCCGCTTCTTCAAGACGCGCCATGCGCGCAAGCGCGTCAACCA 57952
QY 300 GTTGTGTGTCTCCGCGCAAGTGCAGCGCGCGCGCGCTGCTCCCAAGCGCATCGCGCGT 359
DB 57953 GTTGTGTGTCTCCGCGCAAGTGCAGCGCGCGCGCGCTGCTCCCAAGCGCATCGCGCGT 58012
QY 360 GAAGTGTGTCTCCGCGCAAGTGCAGCGCGCGCGCGCTGCTCCCAAGCGCATCGCGCGCA 419
DB 58013 GAAGTGTGTCTCCGCGCAAGTGCAGCGCGCGCGCGCTGCTCCCAAGCGCATCGCGCGCA 58072
QY 420 GCGGCTGCACTCTGTCGCGCGCGCGCGCGCGCGCTGCTCCGCAAGCGCATCGCGCGT 479
DB 58073 GCGGCTGCACTCTGTCGCGCGCGCGCGCGCGCGCTGCTCCGCAAGCGCATCGCGCGT 58132
QY 480 GCGCTGCTGCACTCTGTCGCGCGCGCGCGCGCGCGCTGCTCCCAAGCGCATCGCGCGT 539
DB 58133 GCGCTGCTGCACTCTGTCGCGCGCGCGCGCGCGCGCTGCTCCCAAGCGCATCGCGCGT 58192
QY 540 CCGGCGCGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 599
DB 58193 CCGGCGCGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 58252
QY 600 CAAGGCAACCAAGCGGAGCTGAGAAACGCTACTAG 636
DB 58253 CAAGGCAACCAAGCGGAGCTGAGAAACGCTACTAG 58289

```

```

RESULT 13
AC068807 208135 bp DNA linear ROD 24-UTL-2001
LOCUS Mus musculus 11 BAC RP23-346P7 (Roswell Park Cancer Institute Human
DEFINITION BAC Library) complete sequence.
ACCESSION AC068807
VERSION AC068807.21 GI:15004885
KEYWORDS HTG.
SOURCE Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 208135)
Metzger, M.L., Lewis, L.R., Hume, J., Edwards, C., Harris, C.,
Dederich, D., Thomas, S., Okunom, G., Carlock, C., Garner, T.,
Addison, S., Pace, A., Williams, G., Bonnin, D., Brooks, A., Brown, J.,
Bunay, C., Davis, C., Burkett, C., Chacko, J., Chen, G., Chen, Z.,
Cox, C., C., Ferraguto, D., Forcum-Taney, J., Sill, R.,
Gorell, J.H., Gunaratne, P., Haller, G., Hernandez, J., Hognes, M.,
Hosak, H., Hou, X., Huber, J., Jackson, L., Jia, Y., Kelly, J., Kelly, S.,
Kover, C., Liu, J., Liu, W., Louiegeed, H., Lozada, R.U., Martin, R.,
Massey, E., McLeod, M.P., Mei, G., Moore, S., Morgan, M., Morris, S.,
Neal, D., Nelson, A., Nguyen, R., Nguyen, N., Ogih, M., Parish, B.,
Perez, D., Reiter, D., Say, J., Shen, H., Vasquez, V., Watlington, S.,
Williamson, A., Wrensford, G., Zhou, X., Bouck, J., Hodgson, A.,
Muzny, D.M., Rives, M., Scherer, S., Sodergren, E., Weinstock, G.,
Worley, K. and Gibbs, R.

```

```

TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 208135)
AUTHORS Worley, K.C.
TITLE Direct Submission
JOURNAL Submitted (10-MAY-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 208135)
Worley, K.C.
TITLE Direct Submission
JOURNAL Submitted (24-JUL-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 24, 2001 this sequence version replaced gi:14787157.
Sequencing is completed to a minimum standard of double strand
coverage with a minimum of 2 clones and 2 reads with no ambiguities
or 2 chemistries with a minimum of 2 clones and 3 reads with no
ambiguities. If the sequence quality does not meet this standard,
it will be indicated in the annotation.

```

The repeat regions shown were identified using RepeatMasker by Adrian Smit.

Sequence similarities were identified using Powerblast by Jinghui Zhang.

Exon/Intron boundaries of identified genes were chosen if there were canonical splice junctions that maintained sequence continuity across the splice junctions.

FEATURES
source
Location/Qualifiers
1..208135

```

repeat_region /organism="Mus musculus"
repeat_region /db_xref="taxon:10090"
repeat_region /chromosome="11"
repeat_region /clone="RP23-346P7"
repeat_region /rpt_family="AT_rich"
repeat_region complement(1028..1167)
repeat_region /rpt_family="RSINE1"
repeat_region complement(1227..1322)
repeat_region /rpt_family="PB1D10"
repeat_region 2242..2356
repeat_region /rpt_family="PB1"
repeat_region 2408..2433
repeat_region /rpt_family="B4"
repeat_region 2434..2546
repeat_region /rpt_family="B1_MM"
repeat_region 2580..2689
repeat_region /rpt_family="(TAGA)n"
repeat_region 2704..2730
repeat_region /rpt_family="(CA)n"
repeat_region 2816..3012
repeat_region /rpt_family="B4"
repeat_region 3061..3117
repeat_region /rpt_family="(CA)n"
repeat_region 3641..3701
repeat_region /rpt_family="(TC)n"
repeat_region 4381..4599
repeat_region /rpt_family="B3"
repeat_region 4606..4818
repeat_region /rpt_family="B2"
repeat_region 5029..5068
repeat_region /rpt_family="(CA)n"
repeat_region complement(5075..5241)
repeat_region /rpt_family="B3"
repeat_region complement(6501..6740)
repeat_region /rpt_family="RSINE1"
repeat_region 6926..7005
repeat_region /rpt_family="B3"
repeat_region 8178..8205
repeat_region /rpt_family="(TG)n"
repeat_region complement(9046..9149)

```

repeat_region	/rpt_family="PB1"	9190.9358
repeat_region	/rpt_family="MER2"	9359.9707
repeat_region	/rpt_family="ORR1D"	9708.9801
repeat_region	/rpt_family="MER2"	9802.9831
repeat_region	/rpt_family="(TTTG)n"	9832.9886
repeat_region	/rpt_family="MER2"	10219.10239
repeat_region	/rpt_family="(TGG)n"	12031.10239
repeat_region	/rpt_family="11MC4"	13166.13231
repeat_region	/rpt_family="MIR"	14452.14426
repeat_region	/rpt_family="MIR"	14411.15037
repeat_region	/rpt_family="12"	15543.15794
repeat_region	/rpt_family="B4"	15849.15890
repeat_region	/rpt_family="(CA)n"	15971.16099
repeat_region	/rpt_family="12"	16336.16404
repeat_region	/rpt_family="BCL_MM"	16419.16550
repeat_region	/rpt_family="B1-F"	17412.17518
repeat_region	/rpt_family="B1-F"	17538.17710
repeat_region	/rpt_family="B3A"	18267.18397
repeat_region	/rpt_family="B3A"	18508.18679
repeat_region	/rpt_family="B3"	18935.18982
repeat_region	/rpt_family="B1_MM"	20446.20476
repeat_region	/rpt_family="(TGG)n"	20593.20889
repeat_region	/rpt_family="B4A"	21091.21091
repeat_region	/rpt_family="PB1D10"	22076.22187
repeat_region	/rpt_family="B4A"	22748.22899
repeat_region	/rpt_family="B4A"	23382.23305
repeat_region	/rpt_family="(A)n"	23355.23355
repeat_region	/rpt_family="B2"	23525.23525
repeat_region	/rpt_family="B2"	24185.24455
repeat_region	/rpt_family="(TGG)n"	24921.25202
repeat_region	/rpt_family="1x6"	25346.25346
repeat_region	/rpt_family="B3"	25444.25444
repeat_region	/rpt_family="PB1D10"	26173.26520
repeat_region	/rpt_family="11MD"	

	Query Match	Best Local Similarity	Matches	Conservative	Mismatches	Indels	Gaps	
repeat_region	26434. .26601	/rpt_family="RSINEL"						
repeat_region	26720. .26801	/rpt_family="B4A"						
repeat_region	26802. .26906	/rpt_family="PBID7"						
repeat_region	27067. .27454	/rpt_family="L1_MM"						
repeat_region	27555. .27575	/rpt_family="(A)n"						
repeat_region	27956. .28217	/rpt_family="B4"						
repeat_region	28275. .28459	/rpt_family="L1M4"						
repeat_region	28460. .28615	/rpt_family="B3A"						
repeat_region	complement(28701. .28884)	/rpt_family="B2"						
repeat_region	28937. .29077	/rpt_family="L1M4"						
repeat_region	29241. .29376	/rpt_family="B4A"						
repeat_region	29323. .29410	/rpt_family="PBID10"						
repeat_region	29438. .29691	/rpt_family="B4A"						
repeat_region	29696. .29823	/rpt_family="(CA)n"						
Query Match	67.1%; Score 426.6; DB 10;	Length 208135;						
Best Local Similarity	95.8%; Pred. No. 8.7e-64;							
Matches 438;	Conservative 0;	Mismatches 19;	Indels 0;	Gaps 0;				
QY 180	AGGCAGACCTCCGACCATCCCTTAGACGCCAAAGATGTGTCCAGTACAGCTTCGGGA	239						
DB 57814	AGGCGACAGCCCCCCTCACGCATCCCTCTCTCCGACAGATGTGTCCAGTACAGCTTCGGGA	57873						
QY 240	GCTGACTACACCCGCTTCTTGACACAGAGGCCCATGCGCCGACGCCCAAAGCCGGTCAACCGA	299						
DB 57874	GCTGACTACACCCGCTTCTTGACACAGAGGCCCATGCGCCGACGCCCAAAGCCGGTCAACCGA	57933						
QY 300	GTTGGTGTGCTCCGAGCCAGTCGGGCCCCGCGCGCTGCTGCCAACGCCATCGGAGCGGT	359						
DB 57934	GTTGGTGTGCTCCGAGCCAGTCGGGCCCCGCGCGCTGCTGCCAACGCCATCGGAGCGGT	57993						
QY 360	GAAGTGTGGGCCCGCCGAGCCGAGATTTCGCTGCATCCCGAGTCGTACCGGCGCA	419						
DB 57994	GAAGTGTGGGCCCGCCGAGCCGAGATTTCGCTGCATCCCGAGTCGTACCGGCGCA	58053						
QY 420	GCGGGTGCAGCTGCTGTGTCCCGGGGGGGCGGCGCGCTTCGGGCMAAGTGTCTTGGT	479						
DB 58054	GCGGGTGCAGCTGCTGTGTCCCGGGGGGGCGGCGCGCTTCGGGCMAAGTGTCTTGGT	58113						
QY 480	GGCCTCGTGAAGTGCAGAGCGCTCACCCGGCTTCACAACCAAGTCGAGACTCAAGAACTT	539						
DB 58114	GGCCTCGTGAAGTGCAGAGCGCTCACCCGGCTTCACAACCAAGTCGAGACTCAAGAACTT	58173						
QY 540	CGGGCGGAGACCGGCGGCGCGCAGAAAGGTTCGCAAGCGCGGCGCGCCGCGGAGNC	599						
DB 58174	CGGGCGGAGACCGGCGGCGCGCAGAAAGGTTCGCAAGCGCGGCGCGCCGCGGAGNC	58233						
QY 600	CAAAGCCAACCAAGCGGAGCTTGAGAAAGCTTACTAG	636						
DB 58234	CAAAGCCAACCAAGCGGAGCTTGAGAAAGCTTACTAG	58270						

RESULT 14
 AL591145
 LOCUS
 DEFINITION Mouse DNA sequence from clone RP23-398F7 on chromosome 11, complete sequence.
 ACCESSION AL591145.24 GI:22474402

Db	122	GACCAAAAGACGGCTTCCAGATACAGTGTGGCGGAGCTGACTTCAACCGGTACAGTACCGA	181
Oy	267	CGGCCCATGCCGACGCGCCAAAGCCGGTACCCGAGTTGGTGTGCTTCGCGCCAGTGCGGCC	326
Db	182	TGGGGCTGGTCGCGACGCGCCAAAGCCGGTACCCGAGCTGGTGTGCTCGGAGCAAGTGCGGCC	241
Oy	367	CGCGGGCTGCTGCGCCCAACGCCATGGGCGGGTGTAAGTGGTGGCGCCGAAACGCAACCGCA	366
Db	242	GGCGGGCTGTGCTGCGCCCAACGCCATGTGGCCGGGCAGATGGTGGCGCCCAACGGGGCCCGA	301
Oy	387	TTTCGGTGCATCCCGATCGCTACCGCGCGCAGCGGGGTGCAGCTGTGTGCCCCGGGGG	446
Db	302	CTTCGGCTGCATCCCGGACCCGTCACCGCGCCAGCGGGGTGCAGCTGTGTGTCTTGGCGG	361
Oy	447	CGCGGGCGCGGCTGTGGCCAAAGTGCCTGTGTGGCCCTCGTGCAAAGTGCAAAGCGCCCTCAC	506
Db	362	CGCGGGCGCGGCGCGGCCCAAGTGCCTGTGTGGCCCTCGTGCAAAGTGCAAAGCGCCCTCAC	421
Oy	507	CCGCTTCCACAACCAAGTGGAGACTCAAGACCTTCGGGCGCGAGACCGCGCGCCGCGCAGAA	566
Db	422	TGCGTTCCACAACCAAGTCCGAGCTCAAGACCTTCGGGCGCGAGACCGCGCGCCGCGCAAC	481
Oy	567	GGGTGCAAAGCCGCGGCGCGCGCGCGCGCGCGCGGAGCCAAAGCCAAACCAAGCGCGA	617
Db	482	GGGCGCGAAGCTGCGGGCCCCGCGCGCGCGCGCGCGCAACCAAGCCAGCGCGGCGCA	532

Search completed: March 29, 2003, 00:43:58
Job time : 2293.01 secs

OY	553	TTTCACAACAGTCCGAGCTCAAGACTTCGGAGCCGAGCGCCTCG-CGCAGAAGGGC	612
Db	554	TTTCCAACAACGTCCGAAGCTCAAGACTTCGGAGCCGAGCGCCTCG-CGCAGAAGGGT	613
OY	613	CGAAGCCCCGCGCCCCGCCGCCGAGCGCCAAAGCCAACCAAGCCAGCTGGAGAAAGCC	672
Db	614	CCCAAGCCCCGCGCCCCGCCGCCGAGCGCCAAAGCCAACCAAGCCAGCTGGAGAAAGCC	673
OY	673	TACTGAGCGCGCGCGCGCCCTCCCACC	702
Db	674	TACTGAGCGAGCGCGCGCTATGAGCC	703
RESULT 2			
LOCUS	BB636457	667 bp	mRNA linear EST 26-OCT-2001
DEFINITION	BB636457 RIKEN full-length enriched, adult male scorta and vein Mus musculus cDNA clone A530001M15 5', mRNA sequence.		
ACCESSION	BB636457		
VERSION	BB636457.1	GI:16472326	
KEYWORDS	EST.		
SOURCE	Mus musculus house mouse.		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.		
REFERENCE	1 (bases 1 to 667)		
AUTHORS	Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Komoto,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saio,R., Sakai,C., Sakai,K., Sano,H., Sasaki,T., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.		
TITLE	RIKEN Mouse ESTs (Arakawa,T., et al. 2001)		
JOURNAL	Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel.: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gscc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/ Carninci,P., Shibata,Y., Hayatani,N., Sugahara,Y., Shibata,K., Itoh,M., Komoto,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) wagi,K., Fujisake,S., Inoue,K., Togawa,Y., Iawa,M., Ohara,E., Watanuki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsunaga,S., Kawai,U., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Kono,H., Fukuishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamana,I., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and Hayashizaki,Y. Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001) Please visit our web site (http://genome.gsc.riken.go.jp) for further details. e mouse tissues.		
FEATURES			
source	location/Qualifiers		
	1..667		
	/organism="Mus musculus"		
	/db_xref="taxon:10090"		

BASE COUNT	ORIGIN
119 a	234 c 212 g 100 t 2 others

```

/clone="A530001M15"
/clone_lib="RIKEN full-length enriched, adult male aorta
and vein"
/sex="male"
/tissue_type="aorta and vein"
/dev_stage="adult"
/lab_host="DH10B"
/notes="Site 1: Sali; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5',
GAGGAGAGAGAGATCCAGACGCTTTTTTTTTTTTTTTVN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 20.0 and subtraction to Rot = 459.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5' GAGGAGAGAGATTCGAGTTAATTAATATATCCCCCCCCCCC
3']. cDNA was cleaved with XhoI and BamHI. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
FLC I."

```

[illegible]

Oligo-dT track not found. Not 1 site shown in beginning of sequence is likely internal to the message. cDNA Library Preparation: M.B. Soares Lab Clone distribution: clones will be available through Research Genetics (www.resgen.com) The following repetitive elements were found in this cDNA sequence: 97-125, >GC rich#low complexity
Seq primer: M13 Forward
POLYA=No.

FEATURES

source

Location/Qualifiers
1. .419
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-B1-10-0-UI"
/clone_lib="UI-R-B1"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT73D-pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-B1 library is a subtracted library derived from the following tissues: atrium at 16.5 dpc, ventricle at 16.5 dpc, AV canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc, AV canal at 15 dpc, ventricle at 13 dpc, and adult heart. For a detailed description of the library from which this clone was derived, please visit our web site at ratseq.eng.uiowa.edu. The subtraction has been previously described in (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)
TAG_SEQ=None found"
BASE COUNT 56 a 135 c 155 g 73 t
ORIGIN

Query Match 46.4%; Score 351.8; DB 10; Length 419;
Best Local Similarity 90.0%; Pred. No. 2,7e-62;
Matches 377; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

185 TGGAGAACCAAGACCATGAAACCGGCGGAGAACCGAGGCGCGCTCCACACCCCT 244
Db 419 TAGAGAACCAAGACCATGAAACCGGCGGAGAACCGAGGCGCGCGCTCCACACCCCT 360
QY 245 TTGAGAACCAAGACCATGAAACCGGCGGAGAACCGAGGCGCGCTCCACACCCCT 304
Db 359 ATGACCAACCAAGACCATGAAACCGGCGGAGAACCGAGGCGCGCTCCACACCCCT 300
QY 305 CCGATGGCGCGCTGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 364
Db 299 CCGATGGCGCGCTGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 240
QY 365 GCCCGCGCGCTGCTCCCAACCGCATGCGCGCGGCAAGTGTGCGACCTAGTGGC 424
Db 239 GCCCGCGCGCTGCTCCCAACCGCATGCGCGCGGCAAGTGTGCGACCTAGTGGC 180
QY 425 CCGATGGCGCGCTGCTCCCAACCGCATGCGCGCGGCAAGTGTGCGACCTAGTGGC 484
Db 179 CCGATGGCGCGCTGCTCCCAACCGCATGCGCGCGGCAAGTGTGCGACCTAGTGGC 120
QY 485 GTGTGAGGCG 544
Db 119 GCGGCG 60
QY 545 TCACCGCGCTTCACACCAAGTGGAGTCAAGACTTCGGAAGTGGAGTGGAGTGGAGTGG 603
Db 59 TCACCGCGCTTCACACCAAGTGGAGTCAAGACTTCGGAAGTGGAGTGGAGTGGAGTGG 1

RESULT 6
LOCUS AI556282/c 360 bp mRNA linear EST 23-MAR-1999
DEFINITION UI-R-C2p-rk-a-03-0-UI.s1 UI-R-C2p Rattus norvegicus cDNA clone
ACCESSION AI556282
VERSION AI556282.1 GI:4488645
KEYWORDS EST.
SOURCE Norway rat.

ORGANISM

Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
AUTHORS 1 (bases 1 to 360)
TITLE Bonaldo,M.F., Lennon,G. and Soares,M.B.
Normalisation and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: meares@blue.weeg.uiowa.edu
Oligo-dT track not found. Not 1 site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.
Fatlma Bonaldo, Ph.D. Clone distribution: clones will be available
through Research Genetics (www.resgen.com) The following repetitive
elements were found in this cDNA sequence: 97-125,
>GC rich#low complexity
Seq primer: M13 Forward.

FEATURES

source

Location/Qualifiers
1. .360
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-C2p-rk-a-03-0-UI"
/clone_lib="UI-R-C2p"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT73D-pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-C2p library is a subtracted library derived from the UI-R-C1 library, which is a subtracted library derived from the UI-R-C0 library. The UI-R-C0 library consisted of a mixture of individually tagged normalized libraries constructed from rat placenta, adult lung, brain, liver, kidney, heart, spleen, ovary, muscle, 8, 12 and 18-day embryo. The tag is a string of 3-5 nucleotides present between the Not I site and the oligo-dT track which allows identification of the library of origin of a clone within the mixture. The subtracted library (UI-R-C2p) was constructed as follows: PCR amplified cDNA inserts from UI-R-C1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with the UI-R-C1 library in the form of single-stranded circles. The remaining single-stranded circles (subtracted library) was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the UI-R-C2p library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996)"
BASE COUNT 54 a 120 c 131 g 54 t 1 others
ORIGIN

Query Match 40.3%; Score 306.2; DB 9; Length 360;
Best Local Similarity 90.6%; Pred. No. 5,7e-55;
Matches 326; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 244 TTGAGAACCAAGACCATGAAACCGGCGGAGAACCGAGGCGCGCTCCACACCCCT 303
Db 360 TTGAGAACCAAGACCATGAAACCGGCGGAGAACCGAGGCGCGCTCCACACCCCT 301
QY 304 ACCGATGGCGCGCTGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 363
Db 300 ACCGATGGCGCGCTGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 241
QY 364 GCGCGCGCGCGCTGCTCCCAACCGCATGCGCGCGGCAAGTGTGCGACCTAGTGGC 423

Db	240	GGCCCGCGCGGAGTGTCCCAAGCCCATCGAGGCGCGGAAAGTGCTGAGCGCCCGAAGAGGA	180
Qy	424	CCCCACTTCGCGTGCATCCCGGACCGCTACCGCGCGACGCGGTGCAGCTGCTGTGTCCC	483
Db	180	CCCGACTTCGCCCTGCATCCCGGATCGTTACCGCGCGACGCGGAGCTGCAGCTGCTGTGCCCC	121
Qy	484	GGTGATGAGGCGCGCGCGCGGCAAGGTGTGCGCTGTGGGCTGTGTCACAGTGCACAACGCG	543
Db	120	GGCGGCGGCGCGCGCGCTCCGCGAAGGTGTGCTGTGGGCGTGTGTCACAGTGCACAACGCG	61
Qy	544	CTCACCGGCTTCACAACCAATCGAGGTCAAGAGCTTCGCGGACCGAGGCGCGCTCGGCGG	603
Db	60	CTCACCGGCTTCACAACCAATCGAGGTCAAGAGCTTCGAGCTTCGAGCTTCGAGACCGCTCGGCGG	1

RESULT 7	LOCUS	DEFINITION	ACCESSTION
A1113131/c	A1113131	315 bp mRNA linear EST 11-FEB-1999	
	UI-R-C2p-nr-f-06-0-UI-s1	UI-R-C2p Rattus norvegicus cDNA clone	
	UI-R-C2p-nr-f-06-0-UI 3'	mRNA sequence.	
	A1113131		

VERSION	ALL3131.1	GI:351
KEYWORDS	EST.	
SOURCE	Norway rat.	
ORGANISM	Rattus norvegicus	

REFERENCE
1 (bases 1 to 315)
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)

JOURNAL	Genome Res. 6 (9) , 791-806 (1996)
MEDLINE	97044477
COMMENT	Contact: Soares, MB

Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565

email: mscares@blue.wei.org, uiowa.edu
0150-dt track not found, Not a site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.
Fatima Bonaldo, Ph.D. Clone distribution: clones will be available
through Research Genetics The following repetitive elements were
found in this cDNA sequence: 97-125, >GC-richLow_complexity
seq primer: M13 forward.

FEATURES	Location/Qualifiers
source	1. .315

```

1. 315
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-C2p-nr-f-06-0-UI"
/clone_1fb="UI-R-C2p"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies)"
/notes="Vector: pTV12-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-C2p
library is a subtracted library derived from the UI-R-C1
library, which is a subtracted library derived from the
UI-R-C0 library. The UI-R-C0 library consisted of a
mixture of individually tagged normalized libraries
constructed from rat placenta, adult lung, brain, liver,
kidney, heart, spleen, ovary, muscle, 8, 12 and 18-day
embryo. The tag is a string of 3-5 nucleotides present
between the Not I site and the oligo-cT track which allows
identification of the library of origin of a clone within
the mixture. The subtracted library (UI-R-C2p) was
constructed as follows: PCR amplified cDNA inserts from
UI-R-C1 clones from which 3' ESTs had been derived was
used as a driver in a hybridization with the UI-R-C1
library in the form of single-stranded circles. The
remaining single-stranded circles (subtracted library) was

```

purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the *ut-R-C2p* library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6: 791-806, 1996)." 46 a 110 c 116 g 43 t

Query Match	33.78;	Score 255.8;	DB 9;	Length 315;
-------------	--------	--------------	-------	-------------

Best local similarity 88.3%; Pred.NO. 1.2E-42;
Matches 278; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

289 TTACCCGCTACGTACCGATGGGCCGTGCCGACGCGCAAGCCGCTACCGAGTGTG 348

315 TACACCGCTTGTGACCGACGGCCGCTGCCGAGTGCACGCGTTCACCGAGTTGGTG 256

349 TGCATCCGGCCAGTGCAGCCCGGCGCGCTGCTGACCAACGCCATCCGCCGCGGCAAGTG 408

255 TGCTGGGCGCAGTCGGGCCCGCGGGCTCTGCCAAGCCATCGGGCGCGTGAAGTGG 196

409 TGGGACCTAGTGGGGCCGACTTCCGCTGCATCCCGACCGCTACCGCGCGCGCGT 468

195 TGGCGCCCGAAGGACCGACTTCCGCTGCATCCGGATGCTAACCGCGCGCAAGCGGTTG 136

469 CACTGCTGTGTCCCGGTGAGGCGCGCGCGCAAGTGGCCCTG 528

b

135 CAGCTGCTGTGCCCCCGGGCGGGCGGGCGCTCGGCCAAGTTCGTATGGTGCCCTCG 76

529 TGCAGTGCAAAGCGCCCTCACCCGCTTCACAAACCACTGGAGCTCAAGGACTTCGGGACC 588

75 TGCAGTGCAGCGCCCTCAGCCCGTTCAACAACCAAGTCGAGCTCAGGGACTTCGGACCT 16

589 GAGGCCGCTCGGCCG 603

b
15 GAGACCGTGCCCG 1

8
 7
 6
 5
 4
 3
 2
 1

Q171546/c
A0171546
E3E hr
PM
1:pm
CCE 17-00PM

Accession	Definition
NC_013701	553 bp DNA linear
HS_3088_B1_H06_MP_CIT	Approved Human Genomic Sperm Library D H
sanjens genomic clone plate-3088 Col-11 Bow-D	N/A

ACCESSION	ORGANISM	CLONE	IMAGE-2000	COL-11	ROW=1	VIEW	SEQUENCE
AQ171546							
AQ171546	1	CT-3568013					
REPSTON							

KEYWORDS: GSS, human capital

ORGANISM Homo sapiens
Phylum: Chordata. Vertebrata. Euteleostomi.
Class: Mammalia.
Order: Primates.
Family: Hominidae.
Genus: Homo.
Species: sapiens.

REFERENCE

1 (pages 1 to 535)

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

AUTHORS
Mahairas, G. G., Wallace, J. C., Smith, K., Swartzell, S., Holzman, T., Keller, A., Shaver, P., Furlong, J., Young, J., Zhao, S., Adams, M. N.

TITLE Sequence-tagged connectors: A sequence approach to mapping and

Proc Natl Acad Sci U S A 95 (17) 9739-9744 (1998)

MEDLINE 99380589
 Contact: Mahajrag GG, Wallace JC, Hood I.

High Throughput Sequencing Center
University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618

Fax: (206) 616-3887
Email: iwallace@u.washington.edu

Sequence Tagged Connector
Plate: 3088 row: P column: 11

Class: BAC ends
High quality sequence stop: 535.

FEATURES	Location/Qualifiers
source	1. .535

```
/organism="Homo sapiens"
/db xref="taxon:9606"
```

100

```

/clone.lib="Plate=3088 Col=11 Row=P"
/clone.lib="CIT Approved Human Genomic Sperm Library D"
/sex="Male"
/notes="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
E-Coli DH10B"
BASE COUNT      98 a      164 c      139 g      132 t      2 others
ORIGIN

Query Match      33.1%; Score 251.2; DB 17; Length 535;
Best Local Similarity 98.8%; Pred. No. 1e-41;
Matches 253; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TACTGAGAGTGGCGTGCCTCCTCTGCTGTGACATGACGCTCCACATGCGCTGTGT 60
DB 304 TACTGAGAGTGGCGTGCCTCCTCTGCTGTGACATGACGCTCCACATGCGCTGTGT 245
QY 61 CTGCTGCTGCTGTGTACACAGCCTTCCGTGTGATGAGGCGCCAGGGGTGGACGGG 120
DB 244 CTGCTGCTGCTGTGTACACAGCCTTCCGTGTGATGAGGCGCCAGGGGTGGACGGG 185
QY 121 TTCAAGAAATATGCGACGAAATCATCCCGAGCTGCGAAGTACCCCGAGCCTCCACCG 180
DB 184 TTCAAGAAATATGCGACGAAATCATCCCGAGCTGCGAAGTACCCCGAGCCTCCACCG 125
QY 181 GAGCTGAGAAACAACAAGCATGAACCGGCGGAGAACGAGGCGGCTCCCAACAC 240
DB 124 GAGCTGAGAAACAACAAGCATGAACCGGCGGAGAACGAGGCGGCTCCCAACAC 65
QY 241 CCCTTTGAGACCAAG 256
DB 64 ACCTTTGAGACCAAG 49

RESULT 9
BE111224/c      291 bp      mRNA      linear      EST 13-JUN-2000
LOCUS      BE111224
DEFINITION      UI-R-BJ1-auw-b-04-0-UI.s1 UI-R-BJ1 Rattus norvegicus cDNA clone
VERSION      BE111224
KEYWORDS      BE111224.1 GI:8503329
SOURCE      EST.
ORGANISM      Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 291)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
JOURNAL
MEDLINE      97044477
COMMENT      Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
Oligo-dT track not found, Not 1 site shown in beginning of sequence
is likely internal to the message. cDNA library preparation: M.B.
Soares Lab Clone distribution: clones will be available through
Research Genetics (www.resgen.com) The following repetitive
elements were found in this cDNA sequence: 97-125,
>C_richlow_complexity
Seq primer: M13 Forward
POLVA=No.

FEATURES
Source      1..291
Location/Qualifiers
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-BJ1-auw-b-04-0-UI"

```

```

/clone.lib="UI-R-BJ1"
/lab host="DH10B (Life Technologies)"
/notes="Vector: p773D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not 1; Site 2: Eco RI; The UI-R-BJ1
library is a subtracted library derived from the following
tissues: atrium at 16.5 dpc, ventricle at 16.5 dpc, AV
canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc,
AV canal at 15 dpc, ventricle at 13 dpc, and adult heart.
For a detailed description of the library from which this
clone was derived, please visit our web site at
rategen.eng.uiowa.edu. The subtraction has been previously
described in (Bonaldo, Lennon and Soares, Genome Research
6:791-806, 1996)
TAG SEQ=None found"
BASE COUNT      41 a      101 c      109 g      40 t
ORIGIN

Query Match      32.0%; Score 243; DB 10; Length 291;
Best Local Similarity 89.7%; Pred. No. 4.8e-40;
Matches 261; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 313 CGGTGCGGAGCGGCGCAAGCGGTGAGTGTGCTCCGCGCAGTGGCGCGGCG 372
DB 291 CGGTGCGGAGCGGCGCAAGCGGTGAGTGTGCTCCGCGCAGTGGCGCGGCG 232
QY 373 CGCTGTGCTGCGCAAGCGGTGAGTGTGCTCCGCGCAGTGGCGCGGCG 432
DB 231 CGCTGTGCTGCGCAAGCGGTGAGTGTGCTCCGCGCAGTGGCGCGGCG 172
QY 433 CGGTGATCCCGGACCGGTATCCGCGCGGCGGCTGCTGTGTCCCGGTGTGAG 492
DB 171 CGGTGATCCCGGACCGGTATCCGCGCGGCGGCTGCTGTGTCCCGGTGTGAG 112
QY 493 GCGCGCGCGCGCGCAAGTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 552
DB 111 GCGCGCGCGCGCGCAAGTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 52
QY 553 TTCCACAACAGTGTGAGCTTCAAGACTTCCGAGCCGAGCGCGCTCGCGCG 603
DB 51 TTCCACAACAGTGTGAGCTTCAAGACTTCCGAGCCGAGCGCGCGCGCG 1

RESULT 10
BF523030      254 bp      mRNA      linear      EST 11-DEC-2000
LOCUS      BF523030
DEFINITION      UI-R-C2D-rk-a-03-0-UI.r1 UI-R-C2D Rattus norvegicus cDNA clone
VERSION      BF523030
KEYWORDS      BF523030.1 GI:11631045
SOURCE      EST.
ORGANISM      Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 254)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
JOURNAL
MEDLINE      97044477
COMMENT      Contact: Soares, MB
Program for Rat Gene Discovery and Mapping
University of Iowa
451 Eckstein Medical Research Building Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: msoares@blue.weeg.uiowa.edu
cDNA library preparation: M.B. Soares Lab Clone distribution:
clones will be available through Research Genetics (www.resgen.com)
This clone is also available through the I.M.A.G.E. Consortium at
LUMC (info@image.lumc.gov). IMAGE ID= 1793134
Seq primer: M13 Forward.

```


Db 596 TTGAAGCGTGTGCGCAGCCGCGCCAGCAGACAGAGCGGAGAGAGCCAGC 655

QY 637 AGGCCCA 643

Db 656 ATATCCA 662

RESULT 12
AK002396

LOCUS 1612 bp mRNA linear HNC 19-JAN-2002

DEFINITION Mus musculus adult male kidney cDNA, RIKEN full-length enriched library, clone:0610009G22:homolog to HYPOTHETICAL 15.3 KDA PROTEIN (FRAGMENT), full insert sequence.

ACCESSION AK002396

VERSION AK002396.1 GI:12832347

KEYWORDS HNC, CAP trapper.

SOURCE Mus musculus (strain: C57BL/6J) adult male kidney cDNA to mRNA, clone 11b:RIKEN full-length enriched mouse cDNA library

ORGANISM Mus musculus

REFERENCE 1 Mammalia, Eutheria, Rodentia, Sciurognathi, Muridae, Murinae, Mus.

AUTHORS Carninci, P. and Hayashizaki, Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

MEDLINE 99279253

PUBMED 10349636

REFERENCE 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE 20493374

PUBMED 11042159

REFERENCE 3

AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagasaki, S., Sasaki, N., Carninci, P., Komno, H., Akiyama, J., Nishi, K., Katsunai, T., Taahito, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

TITLE RIKEN integrated sequence analysis (RISA) system-184-format sequencing pipeline with 384 multichannel sequencer

JOURNAL Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE 20530913

PUBMED 11076861

REFERENCE 4

AUTHORS Kawai, J., Shingawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, T., Hara, A., Fukunishi, Y., Komno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamakata, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., Kling, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L. M., Stambli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Guscinich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzei, J., Mombaur, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohsaki, S. and Hayashizaki, Y.

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409 (6821), 685-690 (2001)

MEDLINE 21085660

PUBMED 11217851

REFERENCE 5 (bases 1 to 1612)

AUTHORS Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hirooka, T., Hori, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Komno, H., Kondo, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, R., Sakai, K., Sano, H., Sasaki, D., Schriml, L., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, Y., Tejima, Y., Toya, T., Yamamura, T., Yamakata, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

TITLE Direct Submission

JOURNAL Submitted (10-UTU-2000) Yoshinide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAGCGCCGCGCAGCTGATTTTATTTATTAACCCGCC 3'], cDNA was prepared by using treated reverse transcriptase and reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 5.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGAGAGCTCAAGCTCATTTATTTATTAACCCGCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: SstI; 3' end: XhoI. Host: SOLR.

FEATURES

source location/Qualifiers

1..1612

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="FANTOM_DB:0610009G22"

/db_xref="MGD:MGI:1896762"

/db_xref="taxon:10090"

/clone="0610009G22"

/sex="male"

/tissue_type="kidney"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="adult"

70..690

/note="data source: SPT, source key: 09YJ03, evidence: ISS homolog to HYPOTHETICAL 15.3 KDA PROTEIN (FRAGMENT)"

putative

/codon_start=1

/protein_id="PAB22068.1"

/db_xref="GI:12832348"

/db_xref="MGD:MGI:1913292"

/translation="MPPPAHLSIPILCTIMNCAFKNDATETIYSHYKVPAP SSSNTLQANRGSHFSTGLDRSRQVQCREIRSKRTISDQCSISPLKLVAG ECLPLVPLWNIIGGYGTWRSRSGEMCVNDKRTORIQLQCDGSTRYKTVV TACKCKRYTROHNESSHNPSPAKPAQHRERKRKSKSKSLSS"

1593..1598

/note="putative"

polyA_signal

polyA_site

BASE COUNT 435 a 397 c 352 g 428 t

ORIGIN

Query Match 11.1%; Score 84.2; DB 11; Length 1612;

Best Local Similarity 52.7%; Pred. No. 1,46-07;

Matches 288; Conservative 0; Mismatches 238; Indels 21; Gaps 4;

29-MAY-2001; 2001WO-US17478.
01-JUN-2000; 2000US-208550P.
04-AUG-2000; 2000US-223542P.
(AMGE-) AMGEN INC.
Paszy CJ, Gao Y;
WPI; 2002-114325/15.
P-PSDB; ABB07207, ABB07209.
New human and mouse cysteine-knot polypeptide designated as Cloaked-2,
for treating or preventing kidney, heart (e.g. myocardial infarction)
or liver (e.g. hepatitis) diseases -
Claim 1; Fig 1; 170dp; English.
The invention relates to polypeptides comprising a cysteine knot motif
and designated as Cloaked-2, derived from human and mouse. The cloaked-2
polypeptides can be expressed by standard recombinant methodology. The
cloaked-2 polynucleotides are useful in gene therapy and antisense
therapy. The cloaked-2 polypeptides and polynucleotides are useful for
treating, preventing, ameliorating or detecting diseases and disorders of
the kidney (e.g. anemia, hypertension or low blood pressure), heart (e.g.
cardiac hypertrophy, congestive heart failure, myocardial infarction,
arrhythmias, atherosclerosis, hypertension or low blood pressure),
skeletal muscle (e.g. muscular dystrophy or cachexia), placenta (e.g.
congenital abnormalities or miscarriage), liver (e.g. hepatitis or
cirrhosis), pancreas (e.g. diabetes or pancreatitis), thyroid (e.g.
Grave's disease or myxedema) or adrenal cortex (e.g. Cushing's disease
or Addison's disease), homeostasis or metabolic diseases (e.g. obesity,
cancer or myopathies), infections, or autoimmune diseases. Selective
binding agents may be used to modulate the biological activities of
Cloaked-2 polypeptides or to detect Cloaked-2 polypeptide levels in a
sample. Transgenic non-human animals are useful for drug candidate
screening. The present sequence represents a cDNA encoding the human
cloaked-2 polypeptide.
Sequence 759 BP; 125 A; 282 C; 244 G; 108 T; 0 other;

Query Match	100.0%	Score 759	DB 24	Length 759
Best Local Similarity	100.0%	Pred. No. 4.3e-138		
Matches 759	0	Mismatches	0	Gaps 0

QY	1	TACTGGAAGTGGCGGCGCCCTCTCTGGGTGGGATACATCAGCTCCCACTGGCCCTGTGT	60
Db	1	TACTGGAAGTGGCGGCGCCCTCTCTGGGTGGGATACATCAGCTCCCACTGGCCCTGTGT	60
QY	61	CTCGTCTGCTCTGTGTACACACAGCCTTCCGTGTAGTGAAGGGCCAGGGGTGGACAGGG	120
Db	61	CTCGTCTGCTCTGTGTACACACAGCCTTCCGTGTAGTGAAGGGCCAGGGGTGGACAGGG	120
QY	121	TTCAGGAATGATATGCCACGCGAATCATCTCCCGAGCTGGAGAATACCCCGAGCTTCCACGG	180
Db	121	TTCAGGAATGATATGCCACGCGAATCATCTCCCGAGCTGGAGAATACCCCGAGCTTCCACGG	180
QY	181	GAGCTGAGAAATCAACAAGACCATGAATCCGGGCGGAGAACGGAAGGGCGGCTCCCAACAC	240
Db	181	GAGCTGAGAAATCAACAAGACCATGAATCCGGGCGGAGAACGGAAGGGCGGCTCCCAACAC	240
QY	241	CCCTTTGAGACCAAAAGACGTGTCCAGTACAGCTGCCGAGCTGTCACTTCAACCCGCTAC	300
Db	241	CCCTTTGAGACCAAAAGACGTGTCCAGTACAGCTGCCGAGCTGTCACTTCAACCCGCTAC	300
QY	301	GTGACCCGATGGGCGCTGGCCGACAGCCCAAGCCGGTACCCGAGCTGTGTCTCCGCGCAG	360
Db	301	GTGACCCGATGGGCGCTGGCCGACAGCCCAAGCCGGTACCCGAGCTGTGTCTCCGCGCAG	360
QY	361	TGCGGCGCGGGCGCGCTGTGCCCAACGCGCATTCGGCCGGGCAAGTGTGGCGACCTAAT	420
Db	361	TGCGGCGCGGGCGCGCTGTGCCCAACGCGCATTCGGCCGGGCAAGTGTGGCGACCTAAT	420

OY	421	GGCCCCGACCTTCGCGCTGATATCCCGGACCGCTACCGGCGCGGAGCGCGTGGACAGCTGCTGT	480
Db	421	GGGCCCCGACCTTCGCGCTGCATCCCGACCGCTACCGGCGCGGAGCGCGTGGACAGCTGCTGT	480
OY	481	CCCGGTGGTGAAGGCGCGCGCGCGCAAGTGGCGCTGGTGGCGCTGTGCAAGTGCAG	540
Db	481	CCCGGTGGTGAAGGCGCGCGCGCGCAAGTGGCGCTGGTGGCGCTGTGCAAGTGCAG	540
OY	541	CGCCTCACCCGCTTCACAAACAAGTCGAGCTCAAGACTTCGGGACCGAGCGCGCTCG	600
Db	541	CGCCTCACCCGCTTCACAAACAAGTCGAGCTCAAGACTTCGGGACCGAGCGCGCTCG	600
OY	601	CCGCAAGAAAGCCGGAACCGCGGCCCGCGCGCGCGCAAGCCAAAGCCAAACCAAGCCGAG	660
Db	601	CCGCAAGAAAGCCGGAACCGCGGCCCGCGCGCGCGCAAGCCAAAGCCAAACCAAGCCGAG	660
OY	661	CTGGAAGAACGCTCTACTGAGGCCCGGCCCGCTCTCCCAACCGGCGGGGCGCCCGGCGCT	720
Db	661	CTGGAAGAACGCTCTACTGAGGCCCGGCCCGCTCTCCCAACCGGCGGGGCGCCCGGCGCT	720
OY	721	GAAACCGGCGCCCAATTTCTGTCTCTGAGCGCGTGT	759
Db	721	GAAACCGGCGCCCAATTTCTGTCTCTGAGCGCGTGT	759

RESULT 2
AAA29055
ID AAA29055 standard; cDNA; 2301 BP

AC AAA29055;

DT 12-SEP-2000 (first entry)

DE Human TGF-beta binding protein (BEER) cDNA

KW osteopathic; transforming growth factor-beta; TGF-beta; binding protein;

KW bone mineralization; SS.

OS Homo sapiens.

FH	Key	Location/Qualifiers

FT / *tag= a

PN WO200032773-A1

PD 08-JUN-2000.

PF 24-NOV-1999; 99WO-US27990.

PR 27-NOV-1998; 98US-0110283.

PA (DARW-) DARWIN DISCOVERY LTD.

PI Brunkow ME, Galas DJ, Kovacevich B, Mulligan JT, Paeper BW;

XX	WBT - 2000 433331 / 35
XX	

DR P-PSDB; AAY96429.

PT Nucleic acids (I) encoding a transforming growth factor beta binding

PT osteoporosis and fractures

PS Claim 1; Page 114-115; 162pp; English.

CC This cDNA encodes a human transforming growth factor-beta (TGF-beta,

CC to the chromosome 17q12-21. The cDNA and protein may be used for

CC inappropriate BEER expression. For example, they may be used to treat

disorders associated with decreased TGF-beta BP expression. The cDNA or vectors may be administered to treat diseases by rectifying mutations or deletions in a patient's genome that affect the activity of BEER by expressing inactive proteins or to supplement the patient's own production of BEER polypeptides. The nucleic acids may be used for recombinant production of BEER, gene therapy, antisense therapy, as probes for diagnostic assays and for functional studies. BEER may be used to raise antibodies and for identification of BEER modulators. BEER antagonists may be used to increase bone mineral content for the treatment of disorders such as osteopenia, osteoporosis, fractures and other disorders associated with low mineral content.

Sequence 2301 BP; 568 A; 635 C; 614 G; 484 T; 0 other;

Query Match 100.0%; Score 759; DB 21; Length 2301;
Best Local Similarity 100.0%; Pred. No. 4,4e-138;
Matches 759; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 TACTGGAAGGTGGGGTGGTCCCTCTCTGCTGTGATCCATGCACTCCCATGGCCCTGTGT 60
DB 12 TACTGGAAGGTGGGGTGGTCCCTCTCTGCTGTGATCCATGCACTCCCATGGCCCTGTGT 71
QY 61 CTGCTGCTGCTGCTGTGATCCATGCACTCCCTGTGATGAGGGCCGAGGGCCGAGGCG 120
DB 72 CTGCTGCTGCTGCTGTGATCCATGCACTCCCTGTGATGAGGGCCGAGGGCCGAGGCG 131
QY 121 TTCAAGATGATGACGACGAAATATCCCGAGCTCGAGAGTACCCCGAGCTCCACCG 180
DB 132 TTCAAGATGATGACGACGAAATATCCCGAGCTCGAGAGTACCCCGAGCTCCACCG 191
QY 181 GAGCTGGAGAACCAAGACATGATACCGGGCGAGAACGAGGGCGGCTCCCGACAC 240
DB 192 GAGCTGGAGAACCAAGACATGATACCGGGCGAGAACGAGGGCGGCTCCCGACAC 251
QY 241 CCTTTGAGACCAAAAGAGTGTCCGAGTACAGCTGCGGAGCTGCACTTACCCGCTAC 300
DB 252 CCTTTGAGACCAAAAGAGTGTCCGAGTACAGCTGCGGAGCTGCACTTACCCGCTAC 311
QY 301 GTGACCGATGAGCGCGTGTCCGAGCGCAAGCCGCTACCCGAGCTGTGTGCTCCGCG 360
DB 312 GTGACCGATGAGCGCGTGTCCGAGCGCAAGCCGCTACCCGAGCTGTGTGCTCCGCG 371
QY 361 TTGGGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
DB 372 TTGGGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 431
QY 421 GGGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
DB 432 GGGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 491
QY 481 CCGGCTGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
DB 492 CCGGCTGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 551
QY 541 CGGCTACCGCGCTTCCCAACAGTGGAGCTCAAGGACTTCCGAGCCGAGCGCGCTCGG 600
DB 552 CGGCTACCGCGCTTCCCAACAGTGGAGCTCAAGGACTTCCGAGCCGAGCGCGCTCGG 611
QY 601 CCGCAGAGGGCGGAAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
DB 612 CCGCAGAGGGCGGAAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 671
QY 661 CTGGAAGAGCGCTTACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 720
DB 672 CTGGAAGAGCGCTTACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 731
QY 721 GAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 759
DB 732 GAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 770

```

RESULT 3
AAA29061

ID AAA29061 standard; cDNA; 2301 BP.
XX
XX AAA29061;
XX
XX 12-SEP-2000 (first entry)
XX
XX Mutant human TGF-beta binding protein (BEER) cDNA.
XX
XX osteopathic; transforming growth factor-beta; TGF-beta; binding protein;
XX BEER; chromosome 17q12-21; gene therapy; antisense therapy; fracture;
XX bone mineralization; mutant; sclerosteosis; ss.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX FH 48,119
XX FT CDS
XX FT /*tag= a
XX FT /note= "mutant BEER created by sclerosteosis
XX FT nonsense mutation"
XX PN
XX MO200032773-A1.
XX
XX 08-JUN-2000.
XX
XX 24-NOV-1999; 99WO-US27990.
XX
XX 27-NOV-1998; 98US-0110283.
XX
XX (DARW-) DARWIN DISCOVERY LTD.
XX
XX Brunkow ME, Gales DJ, Kovacevich B, Mulligan JT, Paepel BW;
XX Van Ness J, Winkler DG;
XX
XX WPI: 2000-412321/35.
XX P-PSDB; AAY96435.
XX
XX Nucleic acids (1) encoding a transforming growth factor beta binding
XX PT protein, useful for identifying agents for treating osteopenia,
XX PT osteoporosis and fractures
XX PS
XX Claim 27; Page 117-118; 162pp; English.
XX
XX This cDNA encodes a mutant human transforming growth factor-beta
XX (TGF-beta) binding protein (BEER) produced as a result of a nonsense
XX mutation in the BEER coding sequence (C to T mutation at position 117).
XX This mutation has been linked to sclerosteosis in humans by linkage
XX analysis of an Afrikaner population in South Africa. The BEER gene has
XX been localized to the chromosome 17q12-21. The cDNA and protein may be
XX used for prevention, treatment and diagnosis of diseases associated with
XX inappropriate BEER expression. For example, they may be used to treat
XX disorders associated with decreased TGF-beta BP expression. The cDNA or
XX vectors may be administered to treat diseases by rectifying mutations or
XX deletions in a patient's genome that affect the activity of BEER by
XX expressing inactive proteins or to supplement the patient's own production
XX of BEER polypeptides. The nucleic acids may be used for recombinant
XX production of BEER, gene therapy, antisense therapy, as probes for
XX diagnostic assays and for functional studies. BEER may be used to raise
XX antibodies and for identification of BEER modulators. BEER antagonists
XX may be used to increase bone mineral content for the treatment of
XX disorders such as osteopenia, osteoporosis, fractures and other
XX disorders associated with low mineral content.
XX
XX Sequence 2301 BP; 568 A; 634 C; 614 G; 485 T; 0 other;

Query Match 99.8%; Score 757.4; DB 21; Length 2301;
Best Local Similarity 99.8%; Pred. No. 8.9e-138;
Matches 758; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 TACTGGAAGGTGGGGTGGTCCCTCTCTGCTGTGATCCATGCACTCCCATGGCCCTGTGT 60
DB 12 TACTGGAAGGTGGGGTGGTCCCTCTCTGCTGTGATCCATGCACTCCCATGGCCCTGTGT 71
QY 61 CTGCTGCTGCTGCTGTGATCCATGCACTCCCTGTGATGAGGGCCGAGGGCCGAGGCG 120

```

D	b		72	CTCGTCTGCTGGTGATCAACAGCCTTCGTGTACTGAGAGGGCTTAGGGGTGGAGGGG	131
O	y		121	TTCAAAGAATTGCACGGAATAATCATCCCGAGGCTCCGAGATGCCCGAGCTTCCACC	180
D	b		132	TTCAAAGATGTGTGCACGGAATCATCCCGAGCTCGAAGATGCCCGAGCTTCCACC	191
O	y		181	GAGCTGGAGAACACAAGACCATGAAACCGGGCGAGAAACGAGGGCGGCTCCCCCAC	240
D	b		192	GAGCTGGAGAACACAAGACCATGAAACCGGGCGAGAAACGAGGGCGGCTCCCCCAC	251
O	y		241	CCCTTTGAGACCAAGACGTGTCCGAGTACAGCTGCCGAGCTGTCACTTCAACCCGTA	300
D	b		252	CCCTTTGAGACCAAGACGTGTCCGAGTACAGCTGCCGAGCTGTCACTTCAACCCGTA	311
O	y		301	GTGACCCGATGGGCGGTGGCCGAGGGCCAAAGCCGGTCAACCGAGCTGGGTGTCTCCGG	360
D	b		312	GTGACCCGATGGGCGGTGGCCGAGGGCCAAAGCCGGTCAACCGAGCTGGGTGTCTCCGG	371
O	y		361	TGCGGCCCGGGCGCGCTTGTCCCAACGCCATTCGCGCGCGGCAAGTGTGTGGCACTAGT	420
D	b		372	TGCGGCCCGGGCGCGCTTGTCCCAACGCCATTCGCGCGCGGCAAGTGTGTGGCACTAGT	431
O	y		421	GGGCCCCGACTTTCGCTGATCCCAGACCGCTACCCGCGCGAGCGGCTGTGCTGTGT	480
D	b		432	GGGCCCCGACTTTCGCTGATCCCAGACCGCTACCCGCGCGAGCGGCTGTGCTGTGT	491
O	y		481	CCCGTGTGTGAGAGCGCCCGCGCGGCAAGGTGCGCTGTGTGTGCTGTGCAAGTGCAG	540
D	b		492	CCCGTGTGTGAGAGCGCCCGCGCGGCAAGGTGCGCTGTGTGTGCTGTGCAAGTGCAG	551
O	y		541	CGCCTCACCCGCTTCCACACAGTGGAGCTCAAGGACTTCGGGACCCGAGCCGCTCGG	600
D	b		552	CGCCTCACCCGCTTCCACACAGTGGAGCTCAAGGACTTCGGGACCCGAGCCGCTCGG	611
O	y		601	CCGCAAGAGGGCGGAGCGCGGCGCCCGCGCGCGGACGCGCAAGGCAACAGGCCGAG	660
D	b		612	CCGCAAGAGGGCGGAGCGCGGCGCCCGCGCGCGGACGCGCAAGGCAACAGGCCGAG	671
O	y		661	CTGAGAAACGCTACTAGAGCCCGCGCGCCCTTCCCAACCGGCGGGCGCCCGCCCT	720
D	b		672	CTGAGAAACGCTACTAGAGCCCGCGCGCCCTTCCCAACCGGCGGGCGCCCGCCCT	731
O	y		721	GAAACCGGCGCCACATTTCTGTCTCTGTGGCGGTGTT	759
D	b		732	GAAACCGGCGCCACATTTCTGTCTCTGTGGCGGTGTT	770
RESULT 4					
AAA29062					
ID	AAA29062 standard; cDNA, 2301 BP.				
XX					
XX	AAA29062;				
XX					
DT	12-SEP-2000 (first entry)				
XX					
DE	Human TGF-beta binding protein (BBER) variant P38R cDNA.				
KM	osteopathic; transforming growth factor-beta; TGF-beta; binding protein				
KW	BBER; variant; P38R; gene therapy; antisense therapy; fracture;				
KM	chromosome 17q12-21; bone mineralization; ss.				
OS					
XX	Homo sapiens.				
XX					
Key	Location/Qualifiers				
FH	48..689				
FT	/*tag= a				
FT	/label= BBER variant_P38R				
FT	/product= TGF-beta_binding_protein				
XX					
PN	MO200032773-A1.				
XX					
PD	08-JUN-2000.				

XX 24-NOV-1999; 99WO-US27990.
PF
XX
XX 27-NOV-1998; 98US-0110283.
PR
XX
XX (DARW-) DARWIN DISCOVERY LTD.
PA
XX Brunkow ME, Glas DJ, Kovacevich B, Mulligan JT, Paepfer BW;
XX F1 Van Ness J, Winkler DG;
PI
XX WPI; 2000-412321/35.
XX
XX WPI; 2000-412321/35.
XX
XX P-PSDB; AAY96436#.
DR
XX
XX Nucleic acids (I) encoding a transforming growth factor beta binding
PT protein, useful for identifying agents for treating osteopenia,
PT osteoporosis and fractures
XX
XX
PS Claim 27; Page 120-121; 162pp; English.

CC This CD encodes a variant human transforming growth factor-beta
CC (TGF-beta) binding protein designated BEER P38r. The encoded protein
CC comprises a substitution of arginine for the wild-type proline at
CC residue 38. The CDNA and protein may be used for prevention, treatment
CC and diagnosis of diseases associated with inappropriate BEER expression.
CC For example, they may be used to treat disorders associated with
CC decreased TGF-beta BP expression. The CDNA or vectors may be administered
CC to treat diseases by rectifying mutations or deletions in a patient's
CC genome that affect the activity of BEER by expressing inactive proteins
CC or to supplement the patients own production of BEER polypeptides. The
CC nucleic acids may be used for recombinant production of BEER, gene
CC therapy, antisense therapy, as probes for diagnostic assays and for
CC functional studies. BEER may be used to raise antibodies and for
CC identification of BEER modulators. BEER antagonists may be used to
CC increase bone mineral content for the treatment of disorders such as
CC osteopenia, osteoporosis, fractures and other disorders associated with
CC low mineral content.
CC
CC
SQ Sequence 2301 BP; 568 A; 634 C; 615 G; 484 T; 0 other:

Query Match	99.8%;	Score 757.4;	DB 21;	Length 230;
Best Local Similarity	99.9%;	Pred. No. 8.9e-138;		
Matches 758; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

Qy	1	TACTGGAAGTGGGCGCCCTCCTCGGGCTGTATCAATGCAAGCTCCCACTGGCCCTGTGT	60
Db	12	TACTGGAAGTGGGCGCCCTCCTCGGGCTGTATCAATGCAAGCTCCCACTGGCCCTGTGT	71
Qy	61	CTCGTCTGCTCTGTGTACACACAGCTTTCGTGTAGTGAAGGGCCAGGGGTGGCAGCG	120
Db	72	CTCGTCTGCTCTGTGTACACACAGCTTTCGTGTAGTGAAGGGCCAGGGGTGGCAGCG	133
Qy	121	TTCAAGAAATGATGCCACGAAATCATCCCGAGCTCGAGAGTACCCGAGCTTCCACCG	180
Db	132	TTCAAGAAATGATGCCACGAAATCATCCCGAGCTCGAGAGTACCCGAGCTTCCACCG	191
Qy	181	GAGCTGGAGAACAAACAACCATGAACCGGGCGAGAAACGAGGGCGGCTTCCCAACAC	240
Db	192	GAGCTGGAGAACAAACAACCATGAACCGGGCGAGAAACGAGGGCGGCTTCCCAACAC	251
Qy	241	CCCTTTGAAACCAAGAGCGTGTCCGAGTAAACGTCGCGGAGCTGCACTTACCCCGCTAC	300
Db	252	CCCTTTGAAACCAAGAGCGTGTCCGAGTAAACGCTGCGGAGCTGCACTTACCCCGCTAC	311
Qy	301	GTTGACCGATGGGCGCGTCCGCGACGCCCAAGCCGATCAACGAGCTGTGTGCTCCGGCCAG	360
Db	312	GTTGACCGATGGGCGCGTCCGCGACGCCCAAGCCGATCAACGAGCTGTGTGCTCCGGCCAG	371
Qy	361	TGCGGCGCCGCGCGCGCTGTGTCCCAACCGCAATCGGCTCGCGCAAGTGTGCGACCTTAGT	420
Db	372	TGCGGCGCGCGCGCGCTGTGTCCCAACCGCAATCGGCTCGCGCAAGTGTGCGACCTTAGT	431
Qy	421	GGGCGCGAATTCCGCTGCATCCCGACCGCTAACCGCGCGCAGCGGTGACGTGTGTGT	480


```

Db 312 GTGACCATGGGCGGTGCGGAGCGGCAAGCGGCTGACCGAGCTGTGTCTCCGGCAG 371
Qy 361 TGGCGCGCGCGCGCGCTGCTGCTCCCAAGCCATGCGCGCGCGCAAGTGTGGCACTAGT 420
Db 372 TGGCGCGCGCGCGCGCTGCTGCTCCCAAGCCATGCGCGCGCGCAAGTGTGGCACTAGT 431
Qy 421 GGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db 432 GGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 491
Qy 481 CCGCGTGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
Db 492 CCGCGTGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 551
Qy 541 CCGCGTGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600
Db 552 CCGCGTGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 611
Qy 601 CCGCGAAGGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
Db 612 CCGCGAAGGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 671
Qy 661 CTGGAGAACGCTTACTAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 720
Db 672 CTGGAGAACGCTTACTAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 731
Qy 721 GAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 759
Db 732 GAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 770

```

RESULT 2

```

US-09-449-218D-3
; Sequence 3, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 2301
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-449-218D-3

```

```

Query Match          99.8%; Score 757.4; DB 4; Length 2301;
Best Local Similarity 99.8%; Pred. No. 1e-144;
Matches 758; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 TACTGGAAGTGGCGTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
Db 12 TACTGGAAGTGGCGTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 71
Qy 61 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
Db 72 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 131
Qy 121 TTCAAGATGATGCGCAGGAATCATCCCGAGCTCGAGAGTACCCCGAGCTCCACCG 180
Db 132 TTCAAGATGATGCGCAGGAATCATCCCGAGCTCGAGAGTACCCCGAGCTCCACCG 191

```

```

Qy 181 GAGCTGGAACAACAAGACCATGATACCGGCGGAGAGACGAGGCGCGCTCCACAC 240
Db 192 GAGCTGGAACAACAAGACCATGATACCGGCGGAGAGACGAGGCGCGCTCCACAC 251
Qy 241 CCCTTTGAGACCAAGAGCTGTGAGTACAGCTGCGGCGAGCTGCACTTACCCGCTAC 300
Db 252 CCCTTTGAGACCAAGAGCTGTGAGTACAGCTGCGGCGAGCTGCACTTACCCGCTAC 311
Qy 301 GTGACGATGGGCGGTGCGGAGCGGCGGAGCGGCTGACCGAGCTGTGTGCTCGGCGAG 360
Db 312 GTGACGATGGGCGGTGCGGAGCGGCGGAGCGGCTGACCGAGCTGTGTGCTCGGCGAG 371
Qy 361 TGGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
Db 372 TGGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 431
Qy 421 GGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db 432 GGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 491
Qy 481 CCGGCTGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 540
Db 492 CCGGCTGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 551
Qy 541 CGCTTACCGCGCTTCCACACAGTGGAGCTCAAGACTTGGGAGCGGAGCGCTCGG 600
Db 552 CGCTTACCGCGCTTCCACACAGTGGAGCTCAAGACTTGGGAGCGGAGCGCTCGG 611
Qy 601 CCGCGAAGGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
Db 612 CCGCGAAGGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 671
Qy 661 CTGGAGAACGCTTACTAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 720
Db 672 CTGGAGAACGCTTACTAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 731
Qy 721 GAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 759
Db 732 GAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 770

```

RESULT 3

```

US-09-449-218D-7
; Sequence 7, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepker, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 2301
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-449-218D-7

```

```

Query Match          99.8%; Score 757.4; DB 4; Length 2301;
Best Local Similarity 99.8%; Pred. No. 1e-144;
Matches 758; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 TACTGGAAGTGGCGTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
Db 12 TACTGGAAGTGGCGTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 71

```


Sequence 7, Application 06/05
Patent No. 6395511
GENERAL INFORMATION:
APPLICANT: Brunkow, Mary E.
APPLICANT: Galas, David J.
APPLICANT: Kovacevich, Brian
APPLICANT: Mulligan, John T.
APPLICANT: Paepfer, Bryan W.

```

; Sequence 13, Application US/09
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.

```



```

; TITLE OF INVENTION: BONE MINERALIZATION
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 35828
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(35828)
; OTHER INFORMATION: n = A,T,C or G
US-09-449-218D-17

```

```

Query Match          49.0%; Score 372.2; DB 4; Length 35828;
Best Local Similarity 85.9%; Pred. No. 5.3e-67;
Matches 413; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

```

```

QY 222 AGGCGGCGCTCCCGCACCACCTTTGAGACCAAGAAGTGTCCGAGTACAGTCCGCGA 281
DB 20988 AGGCTGACAGCCCTTCACGACATCTCTCTCCGAGATGTCTCGAGTACAGTCCGCGA 21047
QY 282 GCTGCACTTACACCGCTACGTAACGATGAGCGCGGCGGCGGCGGCGGCGGCGGCGG 341
DB 21048 GCTGCACTTACACCGCTTCTGTAAGACGCGGCGGCGGCGGCGGCGGCGGCGGCGG 21107
QY 342 GCTGCTGTCTCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 401
DB 21108 GTTGTGTGTCTCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 21167
QY 402 CAAGTGTGTGTCTCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 461
DB 21168 GAAGTGTGTGTCTCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 21227
QY 462 GCGGCTGTGTCTCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 521
DB 21228 GCGGCTGTGTCTCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 21287
QY 522 GCGGCTGTGTCTCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 581
DB 21288 GCGGCTGTGTCTCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 21347
QY 582 CGGACCGGAGCGCTCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 641
DB 21348 CGGACCGGAGCGCTCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 21407
QY 642 CAAGCCCAACCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 701
DB 21408 CAAGCCCAACCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 21467
QY 702 C 702
DB 21468 C 21468

```

RESULT 12
US-09-188-930-40
Sequence 40, Application US/09188930A
Patent No. 6150502
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Onrust, Rene
APPLICANT: Murison, James Greg
TITLE OF INVENTION: Compositions Isolated From Skin Cells
FILE REFERENCE: 11000.1011c1
CURRENT APPLICATION NUMBER: US/09/188,930A
NUMBER OF SEQ ID NOS: 348

```

; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 40
; LENGTH: 962
; TYPE: DNA
; ORGANISM: mouse
US-09-188-930-40

```

```

Query Match          10.9%; Score 82.6; DB 3; Length 962;
Best Local Similarity 52.5%; Pred. No. 1.1e-08;
Matches 287; Conservative 0; Mismatches 239; Indels 21; Gaps 4;

```

```

QY 112 TGGCAGCGCTTCAAGATGATGCCAGGAATCATCTCCGAGCTCGAGTACCCCGAG 171
DB 107 TGTGTTGCTTTTAAAAAAGATGCGCACAGAAATCTTTATTCATGTGTGTTAACTCTTC 166
QY 172 CCTCAGCGGAGCTGGAACAACAACGATGACCGGCGGAGAGAGAGGCGGCGCT 231
DB 167 CCGGAC--ACCCAGCAGCAACGACCTGATTAAGCCAGAGATGAGGAGGCGCAT 223
QY 232 CCCACCAACC--CTTGAAGCAAAAGACGTGTCCAGTACAGTCCCGGAGCTGCAC 288
DB 224 TTCACTAGCACTGACCTGATGTAACAGTGAAGTGTGAGGAGGAGGAGGAGTGGG 283
QY 289 TTACCCCGCTACGTGACCGATGAGCGCTGCGGAGCGGCGGCGGCGGCGGCGGCGG 348
DB 284 TCCACCAATATCATTTCCGAGCGGCGGAGTGCACAGCATCAGCCCTGTAAGAGAGCTGTG 343
QY 349 TGTCTCCGCGCAGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 402
DB 344 TCGCGGCGGAGTGTCTTCCCTCTGCGGTGTCTTCCAACTGATGAGAGAGGAGTCA 403
QY 403 -----AAGTGTGTGCGACCTAGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 456
DB 404 ACAAGTACTGTGAGCGGAGAGAGAGCTCTCAGAGTGTGCGGTGTGTAAGAAAGAGCGG 463
QY 457 GCGCAGCGGTGTGAGCTGTGTCTCCGAGTGTGAGCGGCGGCGGCGGCGGCGGCGGCGG 516
DB 464 ACCCAGAGAGTGTGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 520
QY 517 CTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 576
DB 521 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 580
QY 577 GACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 636
DB 581 TTTGAAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 640
QY 637 AGCGCCA 643
DB 641 AATATCA 647

```

RESULT 13
US-09-188-930-209
Sequence 209, Application US/09188930A
Patent No. 6150502
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Onrust, Rene
APPLICANT: Murison, James Greg
TITLE OF INVENTION: Compositions Isolated From Skin Cells
FILE REFERENCE: 11000.1011c1
CURRENT APPLICATION NUMBER: US/09/188,930A
NUMBER OF SEQ ID NOS: 348
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 209
LENGTH: 962
TYPE: DNA
ORGANISM: Mouse

US-09-188-930-209

Query Match 10.9%; Score 82.6; DB 3; Length 962;
 Best Local Similarity 52.5%; Pred. No. 1.1e-08;
 Matches 287; Conservative 0; Mismatches 239; Indels 21; Gaps 4;

112 TGGCAGCGCTTCAAGATGATGACGAGGAATCATCCCGAGCTCGAGATACCCCGG 171
 107 TGTTCGCTTTTAAATGATGACAGGAATCTTTATTCATGATGTTAACTGTTC 166
 172 CTCTCCAGCGAGCTGAGAAACAAGACCATGAAACCGGCGAGAAACGAGGCGGCT 231
 167 CCGGACAC---ACCCAGACAGCAACAGCCCTGATCAACCGAGATGAGCAAGCAT 223
 232 CCCACACACCC---CTTTGAGACCAAGAGATGTCGAGTACAGCTGCGCGAGCTGCAC 288
 224 TTCACTGACACTGAGCTGATTCGAAACATGATTCAGTGGAGCTGCAAGGAACTGCGG 283
 289 TTCAACCGCTACGTCAGCAGATGAGCGCTGCGAGCGCAGCGGTCACCGAGCTGTG 348
 284 TCACCAAAATACATTTTCGAGCGGCACTGACAGCATACGCTTGAAGAGCTGTG 343
 349 TGCTCCGCGAGTGGCGGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 402
 344 TGCGCGGCGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 403
 403 -----AAGTGTGGGAGCCTAGTGGGCGGACCTTCCGCTGCAATCCCGACCGCTACCGG 456
 404 ACAAGTACTGAGAGCGGAGAGAGCTCTCAAGAGTGGGTGTGCAACGACAGACGCGG 463
 457 GCGAGCGCTGAGCTGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 516
 464 ACCAGAGATTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 520
 517 CTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 576
 521 GTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 580
 577 GACTTGGGAGCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 636
 581 TTGGAAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 640
 QY 637 AGCGCCA 643
 DB 641 AATCCA 647

RESULT 14
 US-08-468-847B-1
 ; Sequence 1, Application US/08468847B
 ; Patent No. 5780263
 ; GENERAL INFORMATION:
 ; APPLICANT: Hastings, Gregg A. and Adams, Mark D.
 ; TITLE OF INVENTION: Human CCN-Like Growth Factor
 ; NUMBER OF SEQUENCES: 20
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: CABELLA, BYRNE, BAIN, GILFILLIAN,
 ; STREET: 6 BECKER FARM ROAD
 ; CITY: ROSELAND
 ; STATE: NEW JERSEY
 ; COUNTRY: USA
 ; ZIP: 07068
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5 INCH DISKETTE
 ; COMPUTER: IBM PS/2
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: WORD PERFECT 5.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/468,847B
 ; FILING DATE: 6 June 1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:

APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: MULLINS, J.G.
 REGISTRATION NUMBER: 33,073
 REFERENCE/DOCKET NUMBER: 325800-442
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201-994-1700
 TELEFAX: 201-994-1744
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 900 BASE PAIRS
 TYPE: NUCLEIC ACID
 STRANDEDNESS: SINGLE
 TOPOLOGY: LINEAR
 MOLECULE TYPE: CDNA
 US-08-468-847B-1

Query Match 9.2%; Score 69.6; DB 1; Length 900;
 Best Local Similarity 50.0%; Pred. No. 4.6e-06;
 Matches 239; Conservative 0; Mismatches 224; Indels 15; Gaps 2;

110 GGTGCGAGCGCTTCAAGATGATGACGAGGAATCATCCCGAGCTGAGAGTACCCCG 169
 176 GGTGCTTGTGCTTTTAAATGATGACAGGAATCTTTATTCATGATGTTAACTGT 235
 170 AGCTTCCAGCGGAGCTGAGAAACAAGACCATGAAACCGGCGAGAAACGAGGCGGCG 229
 236 TTCAAGCAGACCCCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 295
 230 CTCCCAACACCCCTTTGAGACCAAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 289
 236 TCAGTAACTGAGACTGATGAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 355
 230 TCACCGCTGATGATGAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 349
 356 CCACCAATAATCATCTGATGAGTGGCAGTGCACACGATCAGGCTCTTAAGAGTGTGT 415
 350 GGTGCGGAGTGGGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 402
 416 GTGTGCGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 475
 403 -----AAGTGTGGGAGCCTAGTGGGCGGCGCTTCCGCTGATCCCGACCGCTACCGCG 457
 476 CAAGTACTGAGAGAGAGAGAGAGCTCCAGAGTGGGCTGTGTAAGTAAACCGGTA 535
 458 GCGAGCGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 517
 536 CCAGAGAGATCCAGCTGAGTGGCCAGATGG---CAGACACAGCAGCTTACAAATTCACAG 592
 518 TGGTGGCTGCTGAGCAAGTGAAGGCGCTCACCGCTTCCCAACCATGCGAGCTCAA 575
 593 TAGTACTGCTGCAAGTGAAGTGAACCCGCGCAGCAGCAGTCCAGTACAA 650

RESULT 15
 US-09-602-877A-92/C
 ; Sequence 92, Application US/09602877A
 ; Patent No. 6432707
 ; GENERAL INFORMATION:
 ; APPLICANT: Reed, Steven G.
 ; APPLICANT: Xu, Jiaochun
 ; APPLICANT: Dillon, David C.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; FILE REFERENCE: 210121.446C5
 ; CURRENT APPLICATION NUMBER: US/09/602,877A
 ; CURRENT FILING DATE: 2000-06-22
 ; NUMBER OF SEQ ID NOS: 107
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 92
 ; LENGTH: 1692
 ; TYPE: DNA

ORGANISM: Homo sapien
US-09-602-877A-92

Query Match 9.2%; Score 69.6; DB 4; Length 1692;
Best Local Similarity 50.0%; Pred. No. 4.7e-06;
Matches 239; Conservative 0; Mismatches 224; Indels 15; Gaps 2;

```
QY 110 GGTGGCAGCGCTTCAAGATGATGCCAGGAATCATCCCGAGCTCGAGAGTACCCCG 169
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1582 GCTGTTTGCTTTTAAAAATGATGCCAGAAATCCTTATTATCATGATGTTTAAACCTG 1523
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 170 AGCCTCACCGGAGCTGGAAGAACAAAGCAATGAACCGGGCGGAGAACGGAGGGGCG 229
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1522 TTCAGACACACCCAGACAGCAAGCAGTTGAATCAAGCCAGAAATGAGGCGGCATT 1463
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 230 CTCCCCACCAACCCCTTTGAGACCAAGACGTGTCGAGTACAGCTGCGGAGCTGCACT 289
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1462 TCAGTAACACTGAGACTGGATCGGAACACTCGGGTTCAAGTGGTTGCCGGAACTGCGTT 1403
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 290 TCACCCGCTACGTGACCGATGGGCCGTGCCGACGCGCAAGCCGGTCAACGAGCTGGTGT 349
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1402 CCACCAATATCATCTGTGATGGCCAGTGCACACAGCATCAGCCCTGTGAMGAGCTGGTGT 1343
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 350 GCTCCCGGCGAGTGGGCCCGGGCGGCTGTGCCCAACGCCATGGGCCCGCC----- 402
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1342 GTGCTGGCGAGTGTGCTTGTCCCTGCAAGTCTCCCTAATGATGAGAGGCTATAGAA 1283
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 403 -----AAGTGTGGCGACCTAGTGGGCCCGACTTCGCTGCATCCCGAACCGCTACCGCG 457
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1282 CAAAGTACTGAGAGAGAGAGAGTCCCAAGAGTGGCCGTGTCAATGACMAAACCCGTA 1223
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 458 CGCAGCGCGTGCAGCTGCTGTGTCCGAGTGAAGGCGCGCGCGCGCAAGGTGGCGC 517
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1222 CCCAGAGAAATCCAGTGCAGTGCACAGATG3---CAGCACACGACCTTCAAAATCACAG 1166
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 518 TGGTGGCTGTGCAAGTGAAGGCGCTCAACCGCTTCCCAACCAAGTTCGAGCTCAA 575
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1165 TAGTCACTGCTGCAAGTGAAGAGTACACCCGCGACACCAAGAGTTCAGTTCACAA 1108
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: March 29, 2003, 01:30:43
Job time : 71.8796 secs

•
•
•
•

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:14:27, Search time 11.6533 Seconds
(without alignments)
479.722 Million cell updates/sec

Title: US-09-867-274-2

Perfect score: 1049
Sequence: 1 OGMQAFKNDATIIPELGEY.....KPRPARSAKNOALENNAY 190

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
4: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
5: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
6: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1049	100.0	213 4 US-09-449-218D-2	Sequence 2, Appl1
2	1040	99.1	213 4 US-09-449-218D-6	Sequence 6, Appl1
3	1040	99.1	213 4 US-09-449-218D-8	Sequence 8, Appl1
4	1033	98.5	213 4 US-09-449-218D-10	Sequence 10, Appl1
5	974	92.9	213 4 US-09-449-218D-14	Sequence 14, Appl1
6	951	90.7	211 4 US-09-449-218D-12	Sequence 12, Appl1
7	901.5	85.9	211 4 US-09-449-218D-16	Sequence 16, Appl1
8	344	32.8	206 4 US-09-188-930-159	Sequence 159, App
9	344	32.8	206 4 US-09-188-930-286	Sequence 286, App
10	341	32.5	206 1 US-08-468-847B-2	Sequence 2, Appl1
11	341	32.5	206 1 US-08-468-847B-20	Sequence 20, Appl1
12	89.5	8.5	184 4 US-09-449-218D-42	Sequence 42, Appl1
13	89.5	8.5	184 4 US-09-440-229B-2	Sequence 2, Appl1
14	87.5	8.3	184 4 US-09-040-229B-6	Sequence 6, Appl1
15	87.5	8.3	735 3 US-09-191-647-9	Sequence 9, Appl1
16	87.5	8.3	735 3 US-09-540-245A-9	Sequence 9, Appl1
17	87.5	8.3	735 4 US-09-540-153-9	Sequence 9, Appl1
18	86.5	8.2	182 4 US-09-040-229B-4	Sequence 4, Appl1
19	86.5	8.2	246 4 US-09-336-093-5	Sequence 5, Appl1
20	86	8.2	184 4 US-09-040-229B-9	Sequence 9, Appl1
21	85	8.1	180 4 US-09-449-218B-44	Sequence 44, Appl1
22	85	8.1	181 4 US-09-040-229B-8	Sequence 8, Appl1
23	81	7.7	270 4 US-08-878-474-1	Sequence 1, Appl1
24	80.5	7.7	267 4 US-09-449-218D-43	Sequence 43, Appl1
25	80	7.6	1480 3 US-09-191-647-7	Sequence 7, Appl1
26	80	7.6	1480 4 US-09-540-245A-7	Sequence 7, Appl1
27	80	7.6	1480 4 US-09-540-153-7	Sequence 7, Appl1

28	80	7.6	1480	4	US-09-182-024A-5	Sequence 5, Appl1
29	80	7.6	1480	5	PCT-US91-09055-2	Sequence 2, Appl1
30	79.5	7.6	830	4	US-09-562-737-39	Sequence 39, Appl1
31	79.5	7.5	336	4	US-08-804-166-8	Sequence 8, Appl1
32	79	7.5	336	4	US-08-910-991-8	Sequence 8, Appl1
33	78	7.4	1210	2	US-08-484-438-7	Sequence 7, Appl1
34	78	7.4	1210	2	US-08-475-035-4	Sequence 4, Appl1
35	77.5	7.4	76	5	PCT-US91-09055-9	Sequence 9, Appl1
36	77	7.3	510	4	US-08-246-489-2	Sequence 2, Appl1
37	74	7.1	470	2	US-08-946-241B-2	Sequence 2, Appl1
38	74	7.1	470	3	US-09-309-053-2	Sequence 2, Appl1
39	74	7.1	479	2	US-08-946-241B-9	Sequence 9, Appl1
40	74	7.1	479	3	US-09-309-053-9	Sequence 9, Appl1
41	73.5	7.0	685	2	US-08-878-989-1	Sequence 1, Appl1
42	73.5	7.0	685	2	US-09-136-282-2	Sequence 2, Appl1
43	73.5	7.0	685	4	US-09-272-796-1	Sequence 1, Appl1
44	73.5	7.0	685	4	US-09-505-744-2	Sequence 2, Appl1
45	73	7.0	301	3	US-08-303-861-21	Sequence 21, Appl1

ALIGNMENTS

```

RESULT 1
US-09-449-218D-2
; Sequence 2, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449, 218D
; CURRENT FILING DATE: 1999-11-24
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-449-218D-2

Query Match          100.0%; Score 1049; DB 4; Length 213;
Best Local Similarity 100.0%; Pred. No. 1.4e-103;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OGMQAFKNDATIIPELGEYPPPELENNKTMAFENGRRPHHPFTKQVSYSCREL 60
DB 24 OGMQAFKNDATIIPELGEYPPPELENNKTMAFENGRRPHHPFTKQVSYSCREL 83
QY 61 HETRYVTGSPGSAKPYVELVCSGCGPARLLPNAIGKMMRPSPGPRCTCJRYRAOR 120
DB 84 HETRYVTGSPGSAKPYVELVCSGCGPARLLPNAIGKMMRPSPGPRCTCJRYRAOR 143
QY 121 VOLLCPGEPAPRARKVRLVASCKKRLTRFNHOSLKDFTGAARPOKGRKPRPARSAK 180
DB 144 VOLLCPGEPAPRARKVRLVASCKKRLTRFNHOSLKDFTGAARPOKGRKPRPARSAK 203
QY 181 ANQAELENNAY 190
DB 204 ANQAELENNAY 213

RESULT 2
US-09-449-218D-6
; Sequence 6, Application US/09449218D
; Patent No. 6395511

```

```
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-449-218D-6
```

```
Query Match          99.1%; Score 1040; DB 4; Length 213;
Best Local Similarity 99.5%; Pred. No. 1.3e-102;
Matches 189; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 QGWAQFNDAATEIIPELGEYPEPPPELENNKTNMRAENGGRPHHPETKDVSEYSCREL 60
DB 24 QGWAQFNDAATEIIPELGEYPEPPPELENNKTNMRAENGGRPHHPETKDVSEYSCREL 83
QY 61 HFTRYVTDGCRSAKPYTELVCSGCGPARLLPNAIGRGKMWPSGDPFCIPDRYRAOR 120
DB 84 HFTRYVTDGCRSAKPYTELVCSGCGPARLLPNAIGRGKMWPSGDPFCIPDRYRAOR 143
QY 121 VOLLCPGGEAPRARKVRLVASCKCKRLTRPHNOSELKDFGTEAPRPOKGRPRPARSAK 180
DB 144 VOLLCPGGEAPRARKVRLVASCKCKRLTRPHNOSELKDFGTEAPRPOKGRPRPARSAK 203
QY 181 ANQAELENAY 190
DB 204 ANQAELENAY 213
```

RESULT 3

```
US-09-449-218D-8
; Sequence 8, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-449-218D-8
```

```
Query Match          99.1%; Score 1040; DB 4; Length 213;
Best Local Similarity 99.5%; Pred. No. 1.3e-102;
Matches 189; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 QGWAQFNDAATEIIPELGEYPEPPPELENNKTNMRAENGGRPHHPETKDVSEYSCREL 60
DB 24 QGWAQFNDAATEIIPELGEYPEPPPELENNKTNMRAENGGRPHHPETKDVSEYSCREL 83
QY 61 HFTRYVTDGCRSAKPYTELVCSGCGPARLLPNAIGRGKMWPSGDPFCIPDRYRAOR 120
DB 84 HFTRYVTDGCRSAKPYTELVCSGCGPARLLPNAIGRGKMWPSGDPFCIPDRYRAOR 143
QY 121 VOLLCPGGEAPRARKVRLVASCKCKRLTRPHNOSELKDFGTEAPRPOKGRPRPARSAK 180
DB 144 VOLLCPGGEAPRARKVRLVASCKCKRLTRPHNOSELKDFGTEAPRPOKGRPRPARSAK 203
QY 181 ANQAELENAY 190
DB 204 ANQAELENAY 213
```

```
DB 24 QGWAQFNDAATEIIPELGEYPEPPPELENNKTNMRAENGGRPHHPETKDVSEYSCREL 83
QY 61 HFTRYVTDGCRSAKPYTELVCSGCGPARLLPNAIGRGKMWPSGDPFCIPDRYRAOR 120
DB 84 HFTRYVTDGCRSAKPYTELVCSGCGPARLLPNAIGRGKMWPSGDPFCIPDRYRAOR 143
QY 121 VOLLCPGGEAPRARKVRLVASCKCKRLTRPHNOSELKDFGTEAPRPOKGRPRPARSAK 180
DB 144 VOLLCPGGEAPRARKVRLVASCKCKRLTRPHNOSELKDFGTEAPRPOKGRPRPARSAK 203
QY 181 ANQAELENAY 190
DB 204 ANQAELENAY 213
```

RESULT 4

```
US-09-449-218D-10
; Sequence 10, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Cercopithecus pygerythrus
US-09-449-218D-10
```

```
Query Match          98.5%; Score 1033; DB 4; Length 213;
Best Local Similarity 98.4%; Pred. No. 6.9e-102;
Matches 187; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 QGWAQFNDAATEIIPELGEYPEPPPELENNKTNMRAENGGRPHHPETKDVSEYSCREL 60
DB 24 QGWAQFNDAATEIIPELGEYPEPPPELENNKTNMRAENGGRPHHPETKDVSEYSCREL 83
QY 61 HFTRYVTDGCRSAKPYTELVCSGCGPARLLPNAIGRGKMWPSGDPFCIPDRYRAOR 120
DB 84 HFTRYVTDGCRSAKPYTELVCSGCGPARLLPNAIGRGKMWPSGDPFCIPDRYRAOR 143
QY 121 VOLLCPGGEAPRARKVRLVASCKCKRLTRPHNOSELKDFGTEAPRPOKGRPRPARSAK 180
DB 144 VOLLCPGGEAPRARKVRLVASCKCKRLTRPHNOSELKDFGTEAPRPOKGRPRPARSAK 203
QY 181 ANQAELENAY 190
DB 204 ANQAELENAY 213
```

RESULT 5

```
US-09-449-218D-14
; Sequence 14, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
US-09-449-218D-14
```

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
FILE REFERENCE: 240083.508
CURRENT APPLICATION NUMBER: US/09/449,218D
NUMBER OF SEQ ID NOS: 45
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 14
LENGTH: 213
TYPE: PRT
ORGANISM: Rattus norvegicus
US-09-449-218D-14

Query Match 92.9%; Score 974; DB 4; Length 213;
Best Local Similarity 92.1%; Pred. No. 1.2e-95;
Matches 175; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 QGWAQKNDATETIIPELGEYEPPEPELENNKTMRANENGRPPHPEPEVXDVSEYSCREL 60
DB 24 QGWAQKNDATETIIPGLREYEPPEPELENNKTMRANENGRPPHHPYDXDVSEYSCREL 83
QY 61 HFRYVTDGCRSAKPTTELVCSCGCGPARLLPNAIGRGKMWPSGDPDRICIDRYRAOR 120
DB 84 HYRFVTDGCRSAKPTTELVCSCGCGPARLLPNAIGRVKMWPNPGDFRICIDRYRAOR 143
QY 121 VOLLCGGAAPRAKRVLVASCKCKRLTRFHNOSLKDFTGAARPOKGRKPRPARSAK 180
DB 144 VOLLCGGAAPRAKRVLVASCKCKRLTRFHNOSLKDFTGAARPOKGRKPRPARSAK 203
QY 181 ANQAELENAY 190
DB 204 ANQAELENAY 213

RESULT 6
US-09-449-218D-12
Sequence 12, Application US/09449218D
Patent No. 6395511
GENERAL INFORMATION:
APPLICANT: Brunkow, Mary E.
APPLICANT: Galas, David J.
APPLICANT: Kovacevich, Brian
APPLICANT: Mulligan, John T.
APPLICANT: Paepert, Bryan W.
APPLICANT: Van Ness, Jeffrey
APPLICANT: Winkler, David G.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
FILE REFERENCE: 240083.508
CURRENT APPLICATION NUMBER: US/09/449,218D
CURRENT FILING DATE: 1999-11-24
NUMBER OF SEQ ID NOS: 45
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 12
LENGTH: 211
TYPE: PRT
ORGANISM: Mus musculus
US-09-449-218D-12

Query Match 90.7%; Score 951; DB 4; Length 211;
Best Local Similarity 89.5%; Pred. No. 3.4e-93;
Matches 170; Conservative 10; Mismatches 8; Indels 2; Gaps 1;

QY 1 QGWAQKNDATETIIPELGEYEPPEPELENNKTMRANENGRPPHPEPEVXDVSEYSCREL 60
DB 24 QGWAQKNDATETIIPGLREYEPPEPELENNKTMRANENGRPPHHPYDXDVSEYSCREL 81
QY 61 HFRYVTDGCRSAKPTTELVCSCGCGPARLLPNAIGRGKMWPSGDPDRICIDRYRAOR 120
DB 82 HYRFVTDGCRSAKPTTELVCSCGCGPARLLPNAIGRVKMWPNPGDFRICIDRYRAOR 141
QY 121 VOLLCGGAAPRAKRVLVASCKCKRLTRFHNOSLKDFTGAARPOKGRKPRPARSAK 180

DB 142 VOLLCGGAAPRAKRVLVASCKCKRLTRFHNOSLKDFTGAARPOKGRKPRPARSAK 201
QY 181 ANQAELENAY 190
DB 202 ANQAELENAY 211

RESULT 7
US-09-449-218D-16
Sequence 16, Application US/09449218D
Patent No. 6395511
GENERAL INFORMATION:
APPLICANT: Brunkow, Mary E.
APPLICANT: Galas, David J.
APPLICANT: Kovacevich, Brian
APPLICANT: Mulligan, John T.
APPLICANT: Paepert, Bryan W.
APPLICANT: Van Ness, Jeffrey
APPLICANT: Winkler, David G.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
FILE REFERENCE: 240083.508
CURRENT APPLICATION NUMBER: US/09/449,218D
CURRENT FILING DATE: 1999-11-24
NUMBER OF SEQ ID NOS: 45
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 16
LENGTH: 176
TYPE: PRT
ORGANISM: Bos taurus
US-09-449-218D-16

Query Match 85.9%; Score 901.5; DB 4; Length 176;
Best Local Similarity 93.8%; Pred. No. 4.7e-88;
Matches 166; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

QY 8 NDATETIIPELGEYEPPEPELENNKTMRANENGRPPHPEPEVXDVSEYSCRELHFTRYVT 67
DB 1 NDATETIIPELGEYEPPEPELENNKTMRANENGRPPHPEPEVXDVSEYSCRELHFTRYVT 59
QY 68 DPCRSKAPVTTELVCSCGCGPARLLPNAIGRGKMWPSGDPDRICIDRYRAORVOLLCRG 127
DB 60 DPCRSKAPVTTELVCSCGCGPARLLPNAIGRGKMWPSGDPDRICIDRYRAORVOLLCRG 119
QY 128 GAAPRAKRVLVASCKCKRLTRFHNOSLKDFTGAARPOKGRKPRPARSAKANQA 184
DB 120 GAAPRAKRVLVASCKCKRLTRFHNOSLKDFTGAARPOKGRKPRPARSAKANQA 176

RESULT 8
US-09-188-930-159
Sequence 159, Application US/09188930A
Patent No. 6150502
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Ornust, Rene
APPLICANT: Murison, James Greg
TITLE OF INVENTION: Compositions and Methods Isolated From Skin Cells
FILE REFERENCE: 11000.1011C1
CURRENT APPLICATION NUMBER: US/09/188,930A
CURRENT FILING DATE: 1998-11-09
NUMBER OF SEQ ID NOS: 348
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 159
LENGTH: 206
TYPE: PRT
ORGANISM: mouse
US-09-188-930-159

Query Match 32.8%; Score 344; DB 4; Length 206;

MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,847B
FILING DATE: 6 June 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: MULLINS, J.G.
REGISTRATION NUMBER: 33,073
REFERENCE/DOCKET NUMBER: 325800-442
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 206 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
US-08-468-847B-20

Query Match 32.5%; Score 341; DB 1; Length 206;
Best Local Similarity 42.7%; Pred. No. 1,5e-28;

Matches 79; Conservative 30; Mismatches 62; Indels 14; Gaps 8;

QY 5 AFNDATTEILPELGEYEP-PPELENNKTMNRAENGGRPPHFFETKDV---EYSCREL 60
DB 23 AFNDATTEIL-VSHVYKPVPAHSSNSTLNQANRGR-HFSNTGLDENTRQVQVCREL 78
QY 61 HFRYTYDGCRCRAKPYTELVCSCGGPARLLPNAIGRG---KWM-RPGCPDFRCIPDRY 116
DB 79 RSKYISDGGCTISPLKELVACGECLPLVPLNMWIGGYGTGYWRRESQEWRCVNDUT 138
QY 117 RAQRVOLLGCGEPAPARKVRLVASCKCKRLTFPHNSELDKDFTEARPOKQRKRPRA 176
DB 139 RTQRIOQCODG-STRTYKITVTACKCKRYTRQHNSHNSPMSDAPVQHHREKKA 197
QY 177 -RSAX 180
DB 198 SKSSX 202

RESULT 12

US-09-449-218D-42

Sequence 42, Application US/09449218D

Patent No. 6395511

GENERAL INFORMATION:

APPLICANT: Brunkow, Mary E.

APPLICANT: Galas, David J.

APPLICANT: Kovacevich, Brian

APPLICANT: Mulligan, John T.

APPLICANT: Paepert, Bryan W.

APPLICANT: Van Ness, Jeffrey

APPLICANT: Winkler, David G.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING

FILE REFERENCE: 240083.508

CURRENT APPLICATION NUMBER: US/09/449,218D

CURRENT FILING DATE: 1999-11-24

NUMBER OF SEQ ID NOS: 45

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 42

LENGTH: 184

TYPE: PRP

ORGANISM: Homo sapiens

US-09-449-218D-42

Query Match 8.5%; Score 89.5; DB 4; Length 184;
Best Local Similarity 24.1%; Pred. No. 0.06;
Matches 38; Conservative 22; Mismatches 57; Indels 41; Gaps 9;

QY 18 GEYEPPELENNK-----TNRAENGGRPPHFFETKDVSYSCRELHFT--R 64
DB 31 GAIP-PDKAQNDSBOTQSPQPGSRNRGRGQGTAMPEE-EVLESSOBALHVTBRK 87
QY 65 YTDGPCR-----SAKPYTELVCSCGGPARLLPNAIGRKWMRPSGDPFR- 110
DB 88 YLRDWCKTQPLKQTHIEBGCNSRTIIRFCYGQCN-SFYIPRIH-----RKIEGSFOS 140
QY 111 ---CIPDRYRAQVOLLGCGEPAP-PAKRVRLVASCKK 144
DB 141 CSFCKPKKFTTMVTLNCPDLQPPTKKKRVTRVQKRC 178

RESULT 13

US-09-040-229B-2

Sequence 2, Application US/09040229B

Patent No. 6432410

GENERAL INFORMATION:

APPLICANT: Harland, Richard

APPLICANT: Hsu, David

TITLE OF INVENTION: Morphogenic Proteins

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP

STREET: 75 DENISE DRIVE

CITY: HILLSBOROUGH

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94010

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/040,229B

FILING DATE: 13-Mar-1998

CLASSIFICATION: <unknown>

ATTORNEY/AGENT INFORMATION:

NAME: OSMAN, RICHARD A

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: 897-020-2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 343-4341

TELEFAX: (650) 343-4342

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 184 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-040-229B-2

Query Match 8.5%; Score 89.5; DB 4; Length 184;

Best Local Similarity 24.1%; Pred. No. 0.06;

Matches 38; Conservative 22; Mismatches 57; Indels 41; Gaps 9;

QY 18 GEYEPPELENNK-----TNRAENGGRPPHFFETKDVSYSCRELHFT--R 64
DB 31 GAIP-PDKAQNDSBOTQSPQPGSRNRGRGQGTAMPEE-EVLESSOBALHVTBRK 87
QY 65 YTDGPCR-----SAKPYTELVCSCGGPARLLPNAIGRKWMRPSGDPFR- 110
DB 88 YLRDWCKTQPLKQTHIEBGCNSRTIIRFCYGQCN-SFYIPRIH-----RKIEGSFOS 140
QY 111 ---CIPDRYRAQVOLLGCGEPAP-PAKRVRLVASCKK 144
DB 141 CSFCKPKKFTTMVTLNCPDLQPPTKKKRVTRVQKRC 178

RESULT 14

US-09-040-229B-6

Sequence 6, Application US/09040229B

Patent No. 6432410

GENERAL INFORMATION:

APPLICANT: Harland, Richard

Hsu, David

TITLE OF INVENTION: Morphogenic Proteins

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

STREET: 75 DENISE DRIVE

CITY: HILLSBOROUGH

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94010

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/040,229B

FILING DATE: 13-Mar-1998

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: OSMAN, RICHARD A

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: B97-020-2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 343-4341

TELEFAX: (650) 343-4342

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 184 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 6:

US-09-040-229B-6

Query Match 8.3%; Score 87.5; DB 4; Length 184;

Best Local Similarity 22.4%; Pred. No. 0.098;

Matches 34; Conservative 24; Mismatches 55; Indels 39; Gaps 8;

QY 23 PPPLE-----NNKTNRAENGRRPHHPETKDVSEVSCRELAFT--RYVTGDP 70

DB 36 PPPKDDPNDESEMQOQSGSRHRRGKGTSMPE--EVLSSQALHITERKYLKRDW 93

QY 71 CR-----SAKPVTELYCSGCGGPARLLPNAIGRKWRPSGDFR----CIP 113

DB 94 CKTOPLQKTHIEGCSRTIINRRCYGCN--SFYIPRHV-----RKEGSPQSCSFCKP 146

QY 114 DRYRQRYVOLLCPGGEAPRARK-VRLVASCCK 144

DB 147 KKFTTMTVTLNCPLEPPRRKKRITRYKRCRC 178

RESULT 15

US-09-191-647-9

Sequence 9, Application US/09191647

Patent No. 6046015

GENERAL INFORMATION:

APPLICANT: Goodman, Corey

APPLICANT: Kid, Thomas

APPLICANT: Brosse, Katja

APPLICANT: Tessier-Lavigne, Marc

TITLE OF INVENTION: Modulating Robo: Ligand Interactions

FILE REFERENCE: B98-031-3

CURRENT APPLICATION NUMBER: US/09/191,647

CURRENT FILING DATE: 1998-11-13

EARLIER APPLICATION NUMBER: 60/065,544

EARLIER FILING DATE: 1997-11-14

EARLIER APPLICATION NUMBER: 60/081,057

EARLIER FILING DATE: 1998-04-07

NUMBER OF SEQ ID NOS: 14

SOFTWARE: Patent Ver. 2.0

SEQ ID NO 9

LENGTH: 735

TYPE: PRT

ORGANISM: Caenorhabditis elegans

US-09-191-647-9

Query Match 8.3%; Score 87.5; DB 3; Length 735;

Best Local Similarity 20.6%; Pred. No. 0.56;

Matches 21; Conservative 17; Mismatches 45; Indels 19; Gaps 2;

QY 44 HHPETKDVSEVSCRELAFTRYVTDGPRSAKPVTELYCSGCGGPARLLPNAIGRKWR 103

DB 644 HPSGEHCDEKRIKCDKQKRRHHIENECRSVDRIKIAECNGYCG----- 687

QY 104 PSGPDFRCIPDRYRQRYVOLLCPGGEAPRARKVRLVASCCK 145

DB 688 --GEONCTAVAKKQKVKMICKNG--TTKISTVHIIRQCGCE 726

Search completed: March 28, 2003, 14:20:15

Job time : 13.6533 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 28, 2003, 18:03:21 ; Search time 40.1204 Seconds
(without alignments)
4861.524 Million cell updates/sec

Title: US-09-867-274-3

Perfect score: 636
Sequence: 1 atgcagccctcactagcccc.....agctgaggaagcctcactag 636

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfileseq1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	634.4	99.7	638	US-09-449-218D-11	Sequence 11, Appl
2	557.6	87.7	674	US-09-449-218D-13	Sequence 13, Appl
3	504.8	79.4	642	US-09-449-218D-9	Sequence 9, Appl
4	501.6	78.9	642	US-09-449-218D-45	Sequence 45, Appl
5	501.6	78.9	2301	US-09-449-218D-1	Sequence 1, Appl
6	501.6	78.9	2301	US-09-449-218D-5	Sequence 5, Appl
7	500	78.6	2301	US-09-449-218D-3	Sequence 3, Appl
8	500	78.6	2301	US-09-449-218D-7	Sequence 7, Appl
9	426.6	67.1	35828	US-09-449-218D-17	Sequence 17, Appl
10	415.8	65.4	532	US-09-449-218D-15	Sequence 15, Appl
11	367	57.7	9301	US-09-449-218D-18	Sequence 18, Appl
12	81.6	12.8	962	US-09-188-930-40	Sequence 40, Appl
13	81.6	12.8	962	US-09-188-930-209	Sequence 209, Appl
14	67.2	10.6	900	US-08-468-8478-1	Sequence 1, Appl
15	65.6	10.3	1692	US-09-602-877A-92	Sequence 92, Appl
16	52.6	8.3	71989	US-09-443-501A-2	Sequence 2, Appl
17	48.8	7.7	44377	US-08-804-227C-7	Sequence 7, Appl
18	48.8	7.7	44377	US-08-804-198-1	Sequence 1, Appl
19	44.4	7.0	8257	US-09-484-970B-65	Sequence 65, Appl
20	44.4	6.9	3930	US-09-162-373-2	Sequence 2, Appl
21	44.2	6.7	11220	US-09-467-946-2	Sequence 2, Appl
22	42.8	6.7	36778	US-09-105-537-32	Sequence 32, Appl
23	42.8	6.7	36778	US-09-105-537-5	Sequence 5, Appl
24	42.8	6.7	38506	US-09-320-878-19	Sequence 19, Appl
25	42.4	6.7	2040	US-08-031-538-10	Sequence 10, Appl
26	42.2	6.6	4050	US-09-543-084A-26	Sequence 26, Appl
27	42.2	6.6	4093	US-09-543-084A-28	Sequence 28, Appl

28	42.2	6.6	4101	US-09-543-084A-27	Sequence 27, Appl
29	41.8	6.6	674	US-09-449-218D-13	Sequence 13, Appl
30	41.4	6.5	953	US-09-452-239-45	Sequence 45, Appl
31	41.4	6.5	1820	US-08-173-508-7	Sequence 7, Appl
32	41.4	6.5	1821	US-08-265-310-7	Sequence 7, Appl
33	41.4	6.5	1821	US-08-951-742-7	Sequence 7, Appl
34	41	6.4	43280	US-08-804-227C-1	Sequence 1, Appl
35	40.8	6.4	15664	US-08-402-282-3	Sequence 3, Appl
36	40.8	6.4	15664	US-08-508-004-3	Sequence 3, Appl
37	40.8	6.4	15664	US-08-402-066-3	Sequence 3, Appl
38	40.8	6.4	15664	US-08-402-068-3	Sequence 3, Appl
39	40.6	6.4	4098	US-09-268-866-1	Sequence 1, Appl
40	40.4	6.4	20235	US-07-642-734C-3	Sequence 3, Appl
41	40.4	6.4	20235	US-08-439-009A-3	Sequence 3, Appl
42	40.4	6.4	23673	US-09-773-816-1	Sequence 1, Appl
43	40.2	6.3	1506	US-09-158-767-7	Sequence 7, Appl
44	40.2	6.3	1506	US-09-158-767-8	Sequence 8, Appl
45	40.2	6.3	1506	US-09-158-767-9	Sequence 9, Appl

ALIGNMENTS

```
RESULT 1
US-09-449-218D-11
; Sequence 11, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepert, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE OF INVENTION: BONE MINERALIZATION
; FILE REFERENCE: 240083.508
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 638
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-449-218D-11
Query Match          99.7%; Score 634.4; DB 4; Length 638;
Best Local Similarity 99.8%; Pred. No. 4.5e-135;
Matches 635; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGCAGCCCTCAGTACCCCGGCTCTCATCTGCTTATCTTGTGAGGCTTGTGTGT 60
    1 ATGCAGCCCTCAGTACCCCGGCTCTCATCTGCTTATCTTGTGAGGCTTGTGTGT 60
Db 1 ATGCAGCCCTCAGTACCCCGGCTCTCATCTGCTTATCTTGTGAGGCTTGTGTGT 60
QY 61 GTGAGAGGCGGAGGCGGAGGCTTCTGAGATGATGACAGAGGCTTCTCCAGGCTT 120
    61 GTGAGAGGCGGAGGCGGAGGCTTCTGAGATGATGACAGAGGCTTCTCCAGGCTT 120
Db 61 GTGAGAGGCGGAGGCGGAGGCTTCTGAGATGATGACAGAGGCTTCTCCAGGCTT 120
QY 121 GAGAGTACCCCGAGGCTCTCTGAGAACCAACAGACCATGATCCGGGCGAGATGGA 180
    121 GAGAGTACCCCGAGGCTCTCTGAGAACCAACAGACCATGATCCGGGCGAGATGGA 180
Db 121 GAGAGTACCCCGAGGCTCTCTGAGAACCAACAGACCATGATCCGGGCGAGATGGA 180
QY 181 GGCAGACCTTCCCATCATCTCTTATGACGCGCAAGATGTGCGAGTATGCTCCGCGAG 240
    181 GGCAGACCTTCCCATCATCTCTTATGACGCGCAAGATGTGCGAGTATGCTCCGCGAG 240
Db 181 GGCAGACCTTCCCATCATCTCTTATGACGCGCAAGATGTGCGAGTATGCTCCGCGAG 240
QY 241 CTGCATTAACCCGCTTCTGACAGCGGCGCATGCGGAGCGGCGGATCAGCGAG 300
    241 CTGCATTAACCCGCTTCTGACAGCGGCGCATGCGGAGCGGCGGATCAGCGAG 300
Db 241 CTGCATTAACCCGCTTCTGACAGCGGCGCATGCGGAGCGGCGGATCAGCGAG 300
QY 301 TTGTGTGTCTCGGCGGCGGCGGCGGCGGCGGCTGTCAGACGCGGCGGCGGCTG 360
```

```
Db 301 TTGGTGTGCTCCGGCCAGTGGGCGCCGGCGCTGCTGCGCAACGCGCATCGGGCGGTG 360
Qy 361 AAGTGTGGGCGCCGGAACGGAATTTCCGTGCATCCCGATCGCTACCGGGCGAG 420
Db 361 AAGTGTGGGCGCCGGAACGGAATTTCCGTGCATCCCGATCGCTACCGGGCGAG 420
Qy 421 CGGTGTGAGCTGTGTGCTCCCGGGGGGCGCGCGCGCTCGCGCAAGTGTCTGTGTG 480
Db 421 CGGTGTGAGCTGTGTGCTCCCGGGGGGCGCGCGCGCTCGCGCAAGTGTCTGTGTG 480
Qy 481 GCTGTGTGCAATGCAAGCGCTTACCGCTTCCACCAAGTGTGAGTCTCAAGACTTC 540
Db 481 GCTGTGTGCAATGCAAGCGCTTACCGCTTCCACCAAGTGTGAGTCTCAAGACTTC 540
Qy 541 GGGCGCGAGACCGCGCGCGCGCAAGAGGTGTGCAAGCGCGCGCGCGCGCGAGGCC 600
Db 541 GGGCGCGAGACCGCGCGCGCGCAAGAGGTGTGCAAGCGCGCGCGCGCGCGAGGCC 600
Qy 601 AAAGCCAAACGAGCGAGCTGTGAGAACGCTTACTAG 636
Db 601 AAAGCCAAACGAGCGAGCTGTGAGAACGCTTACTAG 636
```

```
RESULT 2
US-09-449-218D-13
; Sequence 13, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepel, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE OF INVENTION: BONE MINERALIZATION
; FILE REFERENCE: 240083.508
; CURRENT FILING DATE: 1999-11-24
; CURRENT APPLICATION NUMBER: US/09/449,218D
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 674
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-09-449-218D-13
```

```
Query Match 87.7%; Score 557.6; DB 4; Length 674;
Best Local Similarity 93.0%; Pred. No. 1e-117;
Matches 597; Conservative 0; Mismatches 39; Indels 6; Gaps 1;
Qy 1 ATGCAGCCCTCACTAGCCCGGCTCTCATCTGCTACTTGTGACCGCTGCTTGTGCT 60
Db 33 ATGCAGCTCTCACTAGCCCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 92
Qy 61 GTGAGGCGCGAGGGGTGGCAAGCTTCAAGAAATGATGCCACAGAGTCTATCCGAGGCTT 120
Db 93 GTGAGAGCGCGAGGGGTGGCAAGCTTCAAGAAATGATGCCACAGAAATCATCCCGAGACTC 152
Qy 121 GGAAGTACCCCGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 174
Db 153 AGAGAGTACCCAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 212
Qy 175 AATGAGGCGAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 234
Db 213 AATGAGGCGAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 272
Qy 235 CGGAGGCTGCACTACACCGCTTCTGTGACGAGCGCGCATGCGCGAGCGCGCAAGCGGTC 294
Db 273 CGGAGGCTGCACTACACCGCTTCTGTGACGAGCGCGCATGCGCGAGCGCGCAAGCGGTC 332
```

```
Qy 295 ACCGAGTTGTGTGCTCTCCGGCCAGTGGGCGCCCGCGGCGCTGTGCTGCCCAACGCGCATCGGG 354
Db 333 ACCGAGTTGTGTGCTCTCCGGCCAGTGGGCGCCCGCGGCGCTGTGCTGCCCAACGCGCATCGGG 392
Qy 355 CGCGTGAAGTGTGTGCGCGCGCGCAACGAGCGGATTTCCGCTGATCCCGGATCGCTACCGC 414
Db 393 CGCGTGAAGTGTGTGCGCGCGCGCAACGAGCGGATTTCCGCTGATCCCGGATCGCTACCGC 452
Qy 415 GCGCAGCGGGGTGAGCTGTGTGCTCCCGGGGGGCGCGCGCGCGCGCGCGCTGCGCAAGTGTGCT 474
Db 453 GCGCAGCGGGGTGAGCTGTGTGCTCCCGGGGGGCGCGCGCGCGCGCGCGCTGCGCAAGTGTGCT 512
Qy 475 CTGTGTGCTGTGTGCAATGCAAGCGCTTACCGCTTCCACCAAGTGTGAGTCTCAAG 534
Db 513 CTGTGTGCTGTGTGCAATGCAAGCGCTTACCGCTTCCACCAAGTGTGAGTCTCAAG 572
Qy 535 GACTTGGGCGCGAGAGACCGCGCGCGCGCAAGAGGTGTGCAAGCGCGCGCGCGCGCGCGCG 594
Db 573 GACTTGGGCGCGAGAGACCGCGCGCGCGCAAGAGGTGTGCAAGCGCGCGCGCGCGCGCGCG 632
Qy 595 GGAGCCAAAGCAACCAAGCGAGCTGTGAGAACGCTTACTAG 636
Db 633 GGAGCCAAAGCAACCAAGCGAGCTGTGAGAACGCTTACTAG 674
```

```
RESULT 3
US-09-449-218D-9
; Sequence 9, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepel, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE OF INVENTION: BONE MINERALIZATION
; FILE REFERENCE: 240083.508
; CURRENT FILING DATE: 1999-11-24
; CURRENT APPLICATION NUMBER: US/09/449,218D
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 642
; TYPE: DNA
; ORGANISM: Cercopithecus pygerythrus
US-09-449-218D-9
```

```
Query Match 79.4%; Score 504.8; DB 4; Length 642;
Best Local Similarity 87.9%; Pred. No. 8.6e-106;
Matches 564; Conservative 0; Mismatches 72; Indels 6; Gaps 1;
Qy 1 ATGCAGCCCTCACTAGCCCGGCTCTCATCTGCTACTTGTGACGCTGCTTGTGCT 60
Db 1 ATGCAGCTCTCACTAGCCCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
Qy 61 GTGAGGCGCGAGGGGTGGCAAGCTTCAAGAAATGATGCCACAGAGTCTATCCGAGGCTT 120
Db 61 GTGAGGCGCGAGGGGTGGCAAGCTTCAAGAAATGATGCCACAGAAATCATCCCGAGACTC 120
Qy 121 GGAAGTACCCCGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 174
Db 121 GGAAGTACCCCGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180
Qy 175 AATGAGGCGAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 234
Db 181 AATGAGGCGAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
Qy 235 CGGAGGCTGCACTACACCGCTTCTGTGACGAGCGCGCATGCGCGAGCGCGCAAGCGGTC 294
Db 241 CGGAGGCTGCACTACACCGCTTCTGTGACGAGCGCGCATGCGCGAGCGCGCAAGCGGTC 300
```


QY 295 ACCGAGTTGTTGCTCTCCGCGCAGTGGCCCGCGCGCTGTGCTACCGCCATCGG 354
DB 301 ACCGAGTTGTTGCTCTCCGCGCAGTGGCCCGCGCGCTGTGCTACCGCCATCGG 360
QY 355 CCGGTGAAGTGTGGCGCCCGGAACGACCGGATTTCCGCTGTACCTCCGATCGCTACCG 414
DB 361 CCGGTGAAGTGTGGCGCCCGGAACGACCGGATTTCCGCTGTACCTCCGATCGCTACCG 420
QY 415 GCGCAGCGGTGTGAGCTGTGTGCTCCCGCGCGCGCGCTGTGCTACCGCGATCGCT 474
DB 421 GCGCAGCGGTGTGAGCTGTGTGCTCCCGCGCGCGCGCTGTGCTACCGCGATCGCT 480
QY 475 CTGTGTGCTGTGTGCAAGTGTGCAAGCGCTGTGCTGTGCAAGCGCTGTGCAAGCGCTCAAG 534
DB 481 CTGTGTGCTGTGTGCAAGTGTGCAAGCGCTGTGCTGTGCAAGCGCTGTGCAAGCGCTCAAG 540
QY 535 GACTTGGGCGCGGAGACCGCGCGCGCGCGAGAGGTGTGCAAGCGCTGTGCGCGCGCGG 594
DB 541 GACTTGGGCGCGGAGACCGCGCGCGCGCGAGAGGTGTGCAAGCGCTGTGCGCGCGCGG 600
QY 595 GGAGCCAAAGCCAAACGAGCGCGAGCTGTGAGAAAGCGCTACTAG 636
DB 601 GGAGCCAAAGCCAAACGAGCGCGAGCTGTGAGAAAGCGCTACTAG 642

RESULT 4
US-09-449-218D-45
Sequence 45, Application US/09449218D
Patent No. 6395511
GENERAL INFORMATION:
APPLICANT: Brunkow, Mary E.
APPLICANT: Galas, David J.
APPLICANT: Kovacevich, Brian
APPLICANT: Mulligan, John T.
APPLICANT: Paepel, Bryan W.
APPLICANT: Van Ness, Jeffrey
APPLICANT: Winkler, David G.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
TITLE OF INVENTION: BONE MINERALIZATION
FILE REFERENCE: 240083.508
CURRENT APPLICATION NUMBER: US/09/449,218D
CURRENT FILING DATE: 1999-11-24
NUMBER OF SEQ ID NOS: 45
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 45
LENGTH: 642
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(639)
US-09-449-218D-45

Query Match 78.9%; Score 501.6; DB 4; Length 642;
Best Local Similarity 87.5%; Pred. No. 4.5e-105;
Matches 562; Conservative 0; Mismatches 74; Indels 6; Gaps 1;

QY 1 ATGCAAGCTCTCACTAGCCCGCTGTGCTATGTGCTACTGTGTGCAAGCGCTGTGCT 60
DB 1 ATGCAAGCTCTCACTAGCCCGCTGTGCTATGTGCTACTGTGTGCAAGCGCTGTGCT 60
QY 61 GTGAGAGGCGGAGGTGTGCAAGCGCTGTGCAAGCGCTGTGCAAGCGCTGTGCAAGCGCT 120
DB 61 GTGAGAGGCGGAGGTGTGCAAGCGCTGTGCAAGCGCTGTGCAAGCGCTGTGCAAGCGCT 120
QY 121 GGAGAGTACCCCGAGCTCTCTCC-----TGAGAACCAACGAGCACTGTAACCGGCGGAG 174
DB 121 GGAGAGTACCCCGAGCTCTCTCC-----TGAGAACCAACGAGCACTGTAACCGGCGGAG 180
QY 175 AATGAGAGCAAGCTCTCTCAACATCTCTATGAGCGCAAGAGTGTGTCCAGTACAGCTGC 234
DB 181 AATGAGAGCAAGCTCTCTCAACATCTCTATGAGCGCAAGAGTGTGTCCAGTACAGCTGC 240

QY 225 CCGCAGTGTGCTCTACCGCGCTTCTGTGACAGCGCCCATGCGCGAGCGCAAGCGGTC 294
DB 241 CCGCAGTGTGCTCTACCGCGCTTCTGTGACAGCGCCCATGCGCGAGCGCAAGCGGTC 300
QY 295 ACCGAGTTGTTGCTCTCCGCGCAGTGGCCCGCGCGCTGTGCTACCGCCATCGG 354
DB 301 ACCGAGTTGTTGCTCTCCGCGCAGTGGCCCGCGCGCTGTGCTACCGCCATCGG 360
QY 355 CCGGTGAAGTGTGGCGCCCGGAACGACCGGATTTCCGCTGTACCTCCGATCGCTACCG 414
DB 361 CCGGTGAAGTGTGGCGCCCGGAACGACCGGATTTCCGCTGTACCTCCGATCGCTACCG 420
QY 415 GCGCAGCGGTGTGAGCTGTGTGCTCCCGCGCGCGCGCTGTGCTACCGCGATCGCT 474
DB 421 GCGCAGCGGTGTGAGCTGTGTGCTCCCGCGCGCGCGCTGTGCTACCGCGATCGCT 480
QY 475 CTGTGTGCTGTGTGCAAGTGTGCAAGCGCTGTGCTGTGCAAGCGCTGTGCAAGCGCTCAAG 534
DB 481 CTGTGTGCTGTGTGCAAGTGTGCAAGCGCTGTGCTGTGCAAGCGCTGTGCAAGCGCTCAAG 540
QY 535 GACTTGGGCGCGGAGACCGCGCGCGCGCGAGAGGTGTGCAAGCGCTGTGCGCGCGCGG 594
DB 541 GACTTGGGCGCGGAGACCGCGCGCGCGCGAGAGGTGTGCAAGCGCTGTGCGCGCGCGG 600
QY 595 GGAGCCAAAGCCAAACGAGCGCGAGCTGTGAGAAAGCGCTACTAG 636
DB 601 GGAGCCAAAGCCAAACGAGCGCGAGCTGTGAGAAAGCGCTACTAG 642

RESULT 5
US-09-449-218D-1
Sequence 1, Application US/09449218D
Patent No. 6395511
GENERAL INFORMATION:
APPLICANT: Brunkow, Mary E.
APPLICANT: Galas, David J.
APPLICANT: Kovacevich, Brian
APPLICANT: Mulligan, John T.
APPLICANT: Paepel, Bryan W.
APPLICANT: Van Ness, Jeffrey
APPLICANT: Winkler, David G.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
TITLE OF INVENTION: BONE MINERALIZATION
FILE REFERENCE: 240083.508
CURRENT APPLICATION NUMBER: US/09/449,218D
CURRENT FILING DATE: 1999-11-24
NUMBER OF SEQ ID NOS: 45
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 2301
TYPE: DNA
ORGANISM: Homo sapien
US-09-449-218D-1

Query Match 78.9%; Score 501.6; DB 4; Length 2301;
Best Local Similarity 87.5%; Pred. No. 5.3e-105;
Matches 562; Conservative 0; Mismatches 74; Indels 6; Gaps 1;

QY 1 ATGCAAGCTCTCACTAGCCCGCTGTGCTATGTGCTACTGTGTGCAAGCGCTGTGCT 60
DB 48 ATGCAAGCTCTCACTAGCCCGCTGTGCTATGTGCTACTGTGTGCAAGCGCTGTGCT 107
QY 61 GTGAGAGGCGGAGGTGTGCAAGCGCTGTGCAAGCGCTGTGCAAGCGCTGTGCAAGCGCT 120
DB 108 GTGAGAGGCGGAGGTGTGCAAGCGCTGTGCAAGCGCTGTGCAAGCGCTGTGCAAGCGCT 167
QY 121 GGAGAGTACCCCGAGCTCTCTCC-----TGAGAACCAACGAGCACTGTAACCGGCGGAG 174
DB 168 GGAGAGTACCCCGAGCTCTCTCC-----TGAGAACCAACGAGCACTGTAACCGGCGGAG 227
QY 175 AATGAGAGCAAGCTCTCTCAACATCTCTATGAGCGCAAGAGTGTGTCCAGTACAGCTGC 234

```

Db 228 AACGAGGCGCGCTCCCAACCCCTTTGAGACCAAGACGTGTCCAGTACAGCTGC 287
QY 235 CGGAGAGCTGCACTACCCCGCTTCTGACAGACGCCCAAGCCGAGCCGAGCCGCTC 294
Db 288 CGCAGAGCTGCACTTACCCCGCTTCTGACAGACGCCCAAGCCGAGCCGCTC 347
QY 295 ACCGAGTTGTGTCTCCGAGCAAGTGGCCCGCGCTGTGCTCCCAAGCCATCGAG 354
Db 348 ACCGAGCTGTGTCTCCGAGCAAGTGGCCCGCGCTGTGCTCCCAAGCCATCGAG 407
QY 355 CGCGTGAAGTGTGTCTCCGAGCAAGTGGCCCGCGCTGTGCTCCCAAGCCATCGAG 414
Db 408 CGCGGCAAGTGTGTCTCCGAGCAAGTGGCCCGCGCTGTGCTCCCAAGCCATCGAG 467
QY 415 GCGCAGCGGCTGTGTCTCCGAGCAAGTGGCCCGCGCTGTGCTCCCAAGCCATCGAG 474
Db 468 GCGCAGCGGCTGTGTCTCCGAGCAAGTGGCCCGCGCTGTGCTCCCAAGCCATCGAG 527
QY 475 CTGTGAGCTGTGTCTCCGAGCAAGTGGCCCGCGCTGTGCTCCCAAGCCATCGAG 534
Db 528 CTGTGAGCTGTGTCTCCGAGCAAGTGGCCCGCGCTGTGCTCCCAAGCCATCGAG 587
QY 535 GACTTCGAGCGGAGACCGCGCGCGCGCGAGAGGTGCGCAAGCCGCGCGCGCGCG 594
Db 588 GACTTCGAGCGGAGACCGCGCGCGCGCGAGAGGTGCGCAAGCCGCGCGCGCGCG 647
QY 595 GAGGCCAAGCCACCAAGCGCGAGAGGTGCGCAAGCCGCTTACTAG 636
Db 648 AGCGCCAAAGCCACCAAGCGCGAGAGGTGCGCAAGCCGCTTACTAG 689

```

RESULT 6

```

US-09-449-218D-5
; Sequence 5, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; TITLE OF INVENTION: BONE MINERALIZATION
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 2301
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-449-218D-5

```

```

Query Match 78.9%; Score 501.6; DB 4; Length 2301;
Best Local Similarity 87.5%; Pred. No. 5.3e-105;
Matches 562; Conservative 0; Mismatches 74; Indels 6; Gaps 1;

QY 1 ATGAGCCCTCACTAGCCCGGCTGCTCATCTGCTTGTGACAGCGCTTGTGCT 60
Db 48 ATGAGCTCCCACTGCGCTGTGTCTGTCTGTCTGTGTGTGTGTGTGTGTGT 107
QY 61 GTGAGGCGCGAGGGGTGCAAGCTTCAAGATGATGCCAAGAGTCTATCCAGGGCTT 120
Db 108 GTGAGGCGCGAGGGGTGCAAGCTTCAAGATGATGCCAAGAGTCTATCCAGGGCTT 167
QY 121 GGAAGTACCCCGAGCTTCTTC-----TGAGAACCAACGACCATGAAACCGGGCGAG 174
Db 168 GGAAGTACCCCGAGCTTCTTC-----TGAGAACCAACGACCATGAAACCGGGCGAG 227
QY 175 AATGAGGCAAGACTCCCAATCTCTATGACGCCAAAGATGTGTCCAGTACAGCTGC 234

```

```

Db 228 AACGAGGCGCGCTCCCAACCCCTTTGAGACCAAGACGTGTCCAGTACAGCTGC 287
QY 235 CGGAGAGCTGCACTACCCCGCTTCTGACAGACGCCCAAGCCGAGCCGAGCCGCTC 294
Db 288 CGCAGAGCTGCACTTACCCCGCTTCTGACAGACGCCCAAGCCGAGCCGCTC 347
QY 295 ACCGAGTTGTGTCTCCGAGCAAGTGGCCCGCGCTGTGCTCCCAAGCCATCGAG 354
Db 348 ACCGAGCTGTGTCTCCGAGCAAGTGGCCCGCGCTGTGCTCCCAAGCCATCGAG 407
QY 355 CGCGTGAAGTGTGTCTCCGAGCAAGTGGCCCGCGCTGTGCTCCCAAGCCATCGAG 414
Db 408 CGCGGCAAGTGTGTCTCCGAGCAAGTGGCCCGCGCTGTGCTCCCAAGCCATCGAG 467
QY 415 GCGCAGCGGCTGTGTCTCCGAGCAAGTGGCCCGCGCTGTGCTCCCAAGCCATCGAG 474
Db 468 GCGCAGCGGCTGTGTCTCCGAGCAAGTGGCCCGCGCTGTGCTCCCAAGCCATCGAG 527
QY 475 CTGTGAGCTGTGTCTCCGAGCAAGTGGCCCGCGCTGTGCTCCCAAGCCATCGAG 534
Db 528 CTGTGAGCTGTGTCTCCGAGCAAGTGGCCCGCGCTGTGCTCCCAAGCCATCGAG 587
QY 535 GACTTCGAGCGGAGACCGCGCGCGCGAGAGGTGCGCAAGCCGCGCGCGCGCG 594
Db 588 GACTTCGAGCGGAGACCGCGCGCGCGAGAGGTGCGCAAGCCGCGCGCGCGCG 647
QY 595 GAGGCCAAGCCACCAAGCGCGAGAGGTGCGCAAGCCGCTTACTAG 636
Db 648 AGCGCCAAAGCCACCAAGCGCGAGAGGTGCGCAAGCCGCTTACTAG 689

```

RESULT 7

```

US-09-449-218D-3
; Sequence 3, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; TITLE OF INVENTION: BONE MINERALIZATION
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 2301
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-449-218D-3

```

```

Query Match 78.6%; Score 500; DB 4; Length 2301;
Best Local Similarity 87.4%; Pred. No. 1.2e-104;
Matches 561; Conservative 0; Mismatches 75; Indels 6; Gaps 1;

QY 1 ATGAGCCCTCACTAGCCCGGCTGCTCATCTGCTTGTGACAGCGCTTGTGCT 60
Db 48 ATGAGCTCCCACTGCGCTGTGTCTGTCTGTCTGTGTGTGTGTGTGTGTGT 107
QY 61 GTGAGGCGCGAGGGGTGCAAGCTTCAAGATGATGCCAAGAGTCTATCCAGGGCTT 120
Db 108 GTGAGGCGCGAGGGGTGCAAGCTTCAAGATGATGCCAAGAGTCTATCCAGGGCTT 167
QY 121 GGAAGTACCCCGAGCTTCTTC-----TGAGAACCAACGACCATGAAACCGGGCGAG 174
Db 168 GGAAGTACCCCGAGCTTCTTC-----TGAGAACCAACGACCATGAAACCGGGCGAG 227

```

QY	175	AATGAGGCAACCTCCCAACATCCCTAATGACGCAAAAGATGTCCTCCAGTACAGCTGC	23.4
Db	228	AACGAGGGCGGGCCTCCCAACCACTCTTTGAGACAAAGACGTCTCCAGTACAGCTGC	28.7
QY	235	CGCAGCTGCATACACCCGCTTCTTGAACAGACGGCCATGACCGCAGCTGCACACCGGCTC	29.4
Db	288	CGCAGCTGCACCTTACACCGCTTACGTGACCGATGGGCGTGGCGCAGCCCAACCGCGTTC	34.7
QY	295	AACGAGTTTGGTGTCTCCCGGCAATGGGGGCCCCGGGGGCGTGTGCTGCCAACGGCATTTGGG	35.4
Db	348	ACCGAGCTGGTGTCTCCCGGCACATGGGGGCCCGGCGCGCTGTGCTGCCAACGGCATTTGGC	40.7
QY	355	CGCGTGAAGTGTGGGCGCCCGAAGGACACCGGATTTCCGCTGCATCCCGATCTCACCGC	41.4
Db	408	CGCGGCAAGTGTGTGGCACACTAAGTGGGGCCGACCTTCCGCTGTGACATCCCGCACCGCTACCGC	46.7
QY	415	GCGCAGCGGGGTGCAGCTGTGTGCCCCGGGGGGCGGCGCCGCGCTGTGACGCAAGTGGCGT	47.4
Db	468	GCGCAGCGCGTGCAGCTGTGTGTCTCCGGTGTGTAAGGCGCGCGCGCGCAAGGTGGCGC	52.7
QY	475	CTTGGTGGCTGTGTCAAGTGCACAGCGGCTTCAACCGCTTCCACCAACAGTGGAGCTCAAG	53.4
Db	528	CTTGGTGGCTGTGTCAAGTGCACAGCGGCTTCAACCGCTTCCACCAACAGTGGAGCTCAAG	58.7
QY	535	GACTTTCGGGCGGAGACCGCGCGGCGCGGACGAAGGGGTGCAAGCCGCGGCGCGGCGCCCGG	59.4
Db	588	GACTTTCGGGACCGAGGCGCGCTCGGCGCGGACGAAGGGCGGGAAGCGCGGCGCGCGCGCGG	64.7
QY	595	GGAGCCAAAGCCAAACGAGCGGACCTGAGAGAACCCCTAATCAG	63.6
Db	648	AGCGCCAAAGCCAAACGAGCGGACCGGAGCTGGAGAGAACCTTATCAG	68.9

```

RESULT 8
US-09-449-218D-7
Sequence 7, Application US/09449218D
Patent No. 6395511
GENERAL INFORMATION:
APPLICANT: Brunkow, Mary E.
APPLICANT: Galas, David J.
APPLICANT: Kovacevich, Brian
APPLICANT: Mulligan, John T.
APPLICANT: Paepel, Bryan W.
APPLICANT: Van Ness, Jeffrey
APPLICANT: Winkler, David G.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
FILE REFERENCE: 240083.508
CURRENT APPLICATION NUMBER: US/09/449,218D
CURRENT FILING DATE: 1999-11-24
NUMBER OF SEQ ID NOS: 45
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 7
LENGTH: 2301
TYPE: DNA
ORGANISM: Homo sapien
US-09-449-218D-7

```

Query Match	Similarity	78.6%	Score 500	DB 4	Length 2301
Best Local	Similarity	87.4%	Pred. No. 1.2e-104		
Matches	561	Conservative	0	Mismatches 75	Indels 6
					Gaps 1
Qy	1 ATGCAAGCCCTCACTAGCCCCGTCCTCATCTGCTCACTTGTGACGCTCTCTTGTGCT	60			
Db	48 ATGCAAGCTCCCACTGACCCTGTGTCTCGTCTGCTGCTGTACACACACCTTCGCTGTA	107			
Qy	61 GTGAGAGGCCAGGGGGTGGCAAGCTTCAGGAATGATGCACAGAGGTCAATCCAGGGCTT	120			
Db	108 GTGGAGGGGCCAAGGGGTGGGACAGCGCTTCAAGATGATGCCACGGAATCAATCCGCGAGCTC	167			
Qy	121 GGAGAGTACCCCGAGGCTCTCC-----TGAGAACACAGACCATGAACGGGGCGAG	174			
Db	168 GGAGAGTACCCCGAGGCTCTCACCCGAGCTGGAGAACACAGACCATGAACCGGGCGGAG	227			

QY	175	AATGGAAGGAGACCTCCCAACCATCCCTATGAGCCAAAATATGTGTCGAGTAC	234
Db	228	AACGGAAGGCGGCTTCCCCACCACTTTTGAAGCAGAAACGTGTCCAGTCACTGC	287
QY	235	CGCAGCTGCACCTAACAACCGCTTCTTGACAGACGGCCATGCGCAGCGCAAGCGGTC	294
Db	288	CGGAGCTGCACCTTCAACCGCTACGTGACGATGAGGCGGTGCGCAGCGCAAGCGGTC	347
QY	295	ACCGAGTTGGTGTCTCCGGCCAGTGGGGCCCCGGCGGGCTGCTGGCCACGCCATTCGGG	354
Db	348	ACCCAGCTGGTGTGCTCCGGCCAGTGGGGCGCGCGCTGTGCTCCCAAGCATATCGGC	407
QY	355	CGCGTGAAGTGTGTGGCGCCGAAACGACCGGATTTTCGCTGCATCCCGATCGCTACCGC	414
Db	408	CGCGGCAAGTGTGTGGCACTAATGGGGCCGACCTTCGCTGCATCCCGCAACCGCTACCGC	467
QY	415	GGCGAGCGGGTGCAGCTGCTGTGCCCCCGGGGGCGCGGCCGCGCTTCGCGCAAGTGTGCT	474
Db	468	GGCGAGCGCGTGCAGCTGTGTGTCCGGGTGTGAGGGCGCGCGGCCCAAGGTGTGCGC	527
QY	475	CTGTGGGCTCTGTGTCAGTGTGAAGCGGCTTACCCGCTTCCACAACCAAGTGGAGCTCAAG	534
Db	528	CTGTGGGCTCTGTGTCAGTGTGAAGCGGCTTACCCGCTTCCACAACCAAGTGGAGCTCAAG	587
QY	535	GACTTGGGGCCGAGACCGGCGGGCGCGAGAAAGGTTGCAAGCCGCGGGCCGGCGCCCGG	594
Db	588	GACTTGGGGACCGAGGCGCGCTCGGCCGCAAGAGGGCGGAAAGCGCGGGCCCGCGCCGG	647
QY	595	GGAGCCAAAGCAACCAAGGGGGAGAGCTGGAGAAAGGCTTACTAG	636
Db	648	AGCGCCAAAGCAACCAAGGGCGAGAGCTGGAGAAAGGCTTACTAG	689

```

RESULT 9
US-09-449-218D-17
Sequence 17, Application US/09449218D
Patent No. 6395511
GENERAL INFORMATION:
APPLICANT: Brunkow, Mary E.
APPLICANT: Galae, David J.
APPLICANT: Kovacevich, Brian
APPLICANT: Mulligan, John T.
APPLICANT: Paepfer, Bryan W.
APPLICANT: Van Ness, Jeffrey
APPLICANT: Winkler, David G.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
FILE REFERENCE: 240083.508
CURRENT APPLICATION NUMBER: US/09/449,218D
CURRENT FILING DATE: 1999-11-24
NUMBER OF SEQ ID NOS: 45
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 17
LENGTH: 35828
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)-(35828)
OTHER INFORMATION: n = A,T,C or G
US-09-449-218D-17

```

Query Match	67.1%	Score 426.6	DB 4	Length 35828
Best Local Similarity	95.8%	Pred. No. 6.5e-89		
Matches 438	Conservative 0	Mismatches 19	Indels 0	Gaps 0

QY	180	AGGCAGACCTCCCAACCATCCCTATGACGCCAAGATGTGTCCGATACAGCTTCGCGCA	239
DB	20988	AGGCTGACGCCCTTCACGCATCCCTCTCTCCGAGATGTGTCCGATACAGCTTCGCGCA	21047
QY	240	GCTGCACTACACCCGCTTCTTGACAGAGAGGCCATGCCGACGCCGAAGCCGGTCAACGA	299

Db	21048	GGGAGCTACACACCGCCTTCTGTACAGAGCGGCCCATATCGGACGCGCAACCGGTACCGGA	21107
Qy	300	GTTGCTGTGTCCTCCGACAGTGGCGGCCCGCGCGGCTGTGCTCCAAACGCATTCGGGCGCGT	359
Db	21108	GTTGGTGTGTGCTCCGCGCAGTGGCGGCCCGCGCGGCTGTGCTCCAAACGCATTCGGGCGCGT	21167
Qy	360	GAAAGTGTGTGGCGCCCGGAACCGGAACTTTCCGCTGTGCATCCCGGATTCGCTACACCGGCGCA	419
Db	21168	GAAAGTGTGTGGCGCCCGGAAGCGAACCGGATTTTCGCTGTGCATCCCGGATTCGCTACACCGGCGCA	21227
Qy	420	GCGGGTGCAGCTGCTGTGCGCCCGGGGGCGCGCGCGCTGTGCACCAAGTGTGCTGTGGT	479
Db	21228	GCGGGTGCAGCTGCTGTGTGCCCCGGGGGGCGCGCGCGCTGTGCACCAAGTGTGCTGTGGT	21287
Qy	480	GGCCTTCGTGCAAGTGTCAAGCGGCTTCAACCGGCTTCCAAACAGTGTGGAAGTTCAAGGACTT	539
Db	21288	GGCCTTCGTGCAAGTGTCAAGCGGCTTCAACCGGCTTCCAAACAGTGTGGAAGTTCAAGGACTT	21347
Qy	540	CGGGCCCGGAGACCGCGCGCGGCCCGCGCAAGAGGTGTGCAAGCGCGCGGCCCGCGCGCGCGGAGC	599
Db	21348	CGGGCCCGGAGACCGCGCGCGGCCCGCGCAAGAGGTGTGCAAGCGCGCGGCCCGCGCGCGCGGAGC	21407
Qy	600	CAAGGCAACACGAGCGGAGCTGGAGAACGCTACTAG	636
Db	21408	CAAGGCAACACGAGCGGAGCTGGAGAACGCTACTAG	21444

```

RESULT 10 218D-15
US-09-449-218D-15
; Sequence 15, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepert, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; TITLE OF INVENTION: BONE MINERALIZATION
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 532
; TYPE: DNA
; ORGANISM: Bos taurus
US-09-449-218D-15

```

Query Match	65.4%	Score 415.8	DB 4	Length 532
Best Local Similarity	87.8%	Pred. No. 1.1e-85		
Matches 466; Conservative	0	Mismatches 62	Indels 3	Gaps 1

QY	90	Db	QY	147	Db	QY	207	Db	QY	122	Db	QY	267	Db	QY	327	
GAATATATCCACAGAGGTCATCTCCAGGGCTTGGAGAGTACCCGAGCTCTCTCTGA---	90	GAATATATCCACAGAAATCATCTCCAGACTGGGCGAGTACCCGAGCTCTCTCCAGAGCT	61	GAACAACAGACCATGAAACGGGCGGAGATGAGGCGAGACTTCCCAACATCTCTATGA	206	GACAAACAGACCATGAAACGGGCGGAGAGAGAGAGACTTCCCAACACCCCTTTGA	121	CGCCAAAGATGTGTCCAGTACAGCTGCGCGAGCTGTACACACCCGCTTCTTGACAGA	266	GACCAAAACCGCTTCCGAGTACAGCTTCCGGGAGACTGTGACTTCAACCGTACGTGACCGA	181	CGGCGCATGTCGCGAGCGCGCAAGCGGTACACGAGTGTGTGTGCTCGGAGCAGATGCGGCC	326	TGGGCGCTGTGCGCGAGCGCGCAAGCGGTACAGAGCTGTGTGTGCTCGGAGCAGATGCGGCC	211	CGCGGAGGTGTGTGCGCAAGCGCATCGGGCGGTGAGAGTGTGTGTGCGCGCCCGAACCGGACCGGA	386

Db	242	GGCGGCGCTGTGGCCCAACGCGCATCGGCGCGCGCAAGTGGCGGCCCAAGCGGCGCGA	301
QY	387	TTTTCCGCTGCATCCCCGAGTGGCTACCGCGGCGACGGGATGCACCTGTGTGCCCGGGG	446
Db	302	CTTCGCGTCGATCCCCCAACGCGCTACCGCGGCGACGGGATGCACCTGTGTGTCTTGACGG	361
QY	447	CGCGCGCGCGCGCTCGCGCGCAAGGTGGCTGTGGTGGCTCGTGCAGTGCACAGCGCTCAC	506
Db	362	CGCGCGCGCGCGCGCGCGCAAGGTGGCGCGCTCGTGGCGCTCGTGCAAGTGCACAGCGCTCAC	421
QY	507	CGCGCTTCACAACAGCTGCGAGCTTCAGAGACTTCGCGCGCGAAGACGCGCGCGCGCAGAA	566
Db	422	TGCGTCCACAAACAGTGCAGCTCAAGGACTTCGCGCGCGAAGCGCGCGCGCGCAAC	481
QY	567	GGGTGCGAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCAAGCAACCAACAGCGGGA	617
Db	482	GGCGCGGAAGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCAACAAAGCCAGCGCGGCGCA	532

```

US-09-449-218D-18
RESULT 11
US-09-449-218D-18
; Sequence 18, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepfer, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Wilklert, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 9301
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-449-218D-18

```

Query Match	57.7%;	Score 367;	DB 4;	length 9301;
Best Local Similarity	86.2%;	Pred. No. 1.6e-74;		
Matches 406; Conservative	0;	Mismatches 65;	Indels 0;	Gaps 0;

QY	166	CGGCGGAGAAATGAGGCAAGACTCTCCCAACATCTCTATGAGCGCAAAAGATGTGTCCAG	225
Db	3138	CAGGGGTGGCCAGGCGGCGGCAACCTTCAAGCGCCCTCTTCCACAAGACTGTGCCAG	3197
QY	226	TACAGCTGCCGCGAGGCTGCACTACCCGTTTCGTACAGAGGCGCAATGCCGAGCGCC	285
Db	3198	TACAGCTGCCGCGAGGCTGCACTTACCCGTTCTGTACCGAATGAGGCGCGGCCAGCGCC	3257
QY	286	AAGCGGCTACCGAGTTGTGTGCTCCGGCGAGTGGGCGCCGCGCTGTCGCCAAC	345
Db	3258	AAGCGGCTACCGAGCTGTGTGTCTCCGGCGAGTGGGCGCCGCGCTGTGCCAAC	3317
QY	346	GCCATCGGCGCGCTGTAAGTGTGTGGCGCCCGAAACGAGACCGAATTTCCGTGCATCCCGAT	405
Db	3318	GCCATCGGCGCGCGAAGTGTGTGGCACTTAAGGGGCCGCACTTCCGCTGCATCCCGAC	3377
QY	406	CGCTAACCGCGGAGCGGAGTGCAGCTGTGTGCCCCGGGGGCGCGGCGCGCTCGCGC	465
Db	3378	CGCTAACCGCGGAGCGCGTGCAGCTGTGTGTGCCGGGTGTGAAGCGCGCGCGCGCGC	3437
QY	466	AAGGTGGCTGTGTGTGCTGTGTGCAGAGTGAAGCGCTTACCCGCTTCACAACTCAATGCG	525
Db	3438	AAGGTGGCTGTGTGTGCTGTGTGCAGAGTGAAGCGCTTACCCGCTTCACAACTCAATGCG	3497

QY 526 GAGCTCAAGGACTTTCGGGCGGAGACCGCGCGGCGGAGAGGCTGACAGCCGCGGCC 585
Db 3498 GAGCTCAAGGACTTTCGGGCGGAGACCGCGCGGCTTCGCGGAGAGGCGGAGCCGCGGCC 3557
QY 586 GCGCGCGCGGAGCGCAAGGCGCAACGAGCGGAGCTTGAGAGACCGCTTAC 636
Db 3558 CCGCGCGGAGCGCGCAAGGCGCAACGAGCGGAGCTTGAGAGACCGCTTAC 3608

RESULT 12
US-09-188-930-40
; Sequence 40, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Muriel, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011C1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 40
; LENGTH: 962
; TYPE: DNA
; ORGANISM: mouse
US-09-188-930-40

Query Match 12.8%; Score 81.6; DB 3; Length 962;
Best Local Similarity 52.6%; Pred. No. 3.5e-10;
Matches 283; Conservative 0; Mismatches 234; Indels 21; Gaps 4;

QY 82 GCCTTCAGGATGATGACAGAGGCTATCCAGGCTTGAGAGTACCC---CGAGCT 138
Db 113 GCTTTTAAATGATGACAGAGGCTATCCAGGCTTGAGAGTACCC---CGAGCT 172
QY 139 CCTCTGAGAACCAACGAGCCATGAAACCGGCGGAGAAATGAGGAGACCTTCC---CAC 195
Db 173 CACCCAGAGCAACAGAGCCCTGATCAAGCCAGAGATGAGAGGAGGATTTCACTAGAC 232
QY 196 CATCCCTAGAGCCCAAGATGTGTCCAGTACAGCTGCGGAGGCTGACCTACACCCGC 255
Db 233 ACTGACTGATGAGAAAGTGTGAGTTCAGTGGGCTGAGGAACTGCGGTCCACAA 292
QY 256 TTCTGACAGAGCGGCGCATGCGGAGCGGCAAGCCGCTCACCGAGTGTGTCTCCGC 315
Db 293 TACATTTGAGAGCGGCGCATGAGCAAGCATGAGCCCTCTGAGAGGAGCTGTGTGCGCGGC 352
QY 316 CAGTGGCGCGCGCGGCTGTCTGCGCAAGCCATCGGCGCG---TGAG 363
Db 353 GAGTGTGCGCGCGGCTGTCTGCGCAAGCCATCGGCGCG---TGAG 412
QY 364 TGTGTGCGCGCGGAGGAGCCGATTTCCGCTGATCCCGGATCGCTACCGGCGCGAGCG 423
Db 413 TGGAGCCGAGAGGCTCTTGAAGTGTGGGTGTGTAACGACAGAGCGGCGCCAGAGG 472
QY 424 GTGACAGTGTGTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 483
Db 473 ATCCAGCTGAGTGTGAGAGCGG---AGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 529
QY 484 TGTGCAAGTGAAGCGCTTACCGGCTTCCAGCAACAGTGGAGCTGAAGACTTTCGGG 543
Db 530 GCGTCAAGTGAAGCGGCTTACCGGCTTCCAGCAACAGTGGAGCTGAAGACTTTCGGG 589
QY 544 CCGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 601
Db 590 GTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 647

RESULT 13
US-09-188-930-209
; Sequence 209, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Muriel, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011C1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 209
; LENGTH: 962
; TYPE: DNA
; ORGANISM: Mouse
US-09-188-930-209

Query Match 12.8%; Score 81.6; DB 3; Length 962;
Best Local Similarity 52.6%; Pred. No. 3.5e-10;
Matches 283; Conservative 0; Mismatches 234; Indels 21; Gaps 4;

QY 82 GCCTTCAGGATGATGACAGAGGCTATCCAGGCTTGAGAGTACCC---CGAGCT 138
Db 113 GCTTTTAAATGATGACAGAGGCTATCCAGGCTTGAGAGTACCC---CGAGCT 172
QY 139 CCTCTGAGAACCAACGAGCCATGAAACCGGCGGAGAAATGAGGAGACCTTCC---CAC 195
Db 173 CACCCAGAGCAACAGAGCCCTGATCAAGCCAGAGATGAGAGGAGGATTTCACTAGAC 232
QY 196 CATCCCTAGAGCCCAAGATGTGTCCAGTACAGCTGCGGAGGCTGACCTACACCCGC 255
Db 233 ACTGACTGATGAGAAAGTGTGAGTTCAGTGGGCTGAGGAACTGCGGTCCACAA 292
QY 256 TTCTGACAGAGCGGCGCATGCGGAGCGGCAAGCCGCTCACCGAGTGTGTCTCCGC 315
Db 293 TACATTTGAGAGCGGCGCATGAGCAAGCATGAGCCCTCTGAGAGGAGCTGTGTGCGCGGC 352
QY 316 CAGTGGCGCGCGGCTGTCTGCGCAAGCCATCGGCGCG---TGAG 363
Db 353 GAGTGTGCGCGCGGCTGTCTGCGCAAGCCATCGGCGCG---TGAG 412
QY 364 TGTGTGCGCGCGGAGGAGCCGATTTCCGCTGATCCCGGATCGCTACCGGCGCGAGCGG 423
Db 413 TGGAGCCGAGAGGCTCTTGAAGTGTGGGTGTGTAACGACAGAGCGGCGCCAGAGG 472
QY 424 GTGACAGTGTGTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 483
Db 473 ATCCAGCTGAGTGTGAGAGCGG---AGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 529
QY 484 TGTGCAAGTGAAGCGCTTACCGGCTTCCAGCAACAGTGGAGCTGAAGACTTTCGGG 543
Db 530 GCGTCAAGTGAAGCGGCTTACCGGCTTCCAGCAACAGTGGAGCTGAAGACTTTCGGG 589
QY 544 CCGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 601
Db 590 GTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 647

RESULT 14
US-08-468-847B-1
; Sequence 1, Application US/08468847B
; Patent No. 5780263
; GENERAL INFORMATION:
; APPLICANT: Hastings, Gregg A. and Adams, Mark D.
; TITLE OF INVENTION: Human CCN-Like Growth Factor
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:

```

; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLIAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,847B
; FILING DATE: 6 June 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: MULLINS, J.G.
; REGISTRATION NUMBER: 33,073
; REFERENCE/DOCKET NUMBER: 325800-442
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 900 BASE PAIRS
; TYPE: NUCLEIC ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: CDNA
; US-08-468-847B-1

Query Match      10.6%; Score 67.2; DB 1; Length 900;
Best Local Similarity 50.9%; Pred. No. 6.3e-07;
Matches 274; Conservative 0; Mismatches 243; Indels 21; Gaps 4;

82 GCCTTCAGAGATGATGCGACAGAGGTGATCCCGGCTTGAGAGTACCCCGAGC---CT 138
184 GCTTTTAAATGATGCGACAGAAATCTTTATTCATGTGTTAACTGTTCCAGCA 243
139 CTCTCTGAGAACCAACGACCATGAACCGGCGGAGAGATGAGGACGACCTCCC---CAC 195
244 CACCCGACGACGACGACGACGTTGAATCAAGCAGAAATGAGGACGAGGATTTCACTAAC 303
196 CATCCCTATGACCCCAAGATGTGTCCGAGTACAGCTGCCGCGAGCTGCATACCCGC 255
304 ACTGAGCTGATGGAACACTGCGGTTCAAGTGGTTGCCGGAGACTGGCTTCCACCAA 363
256 TTCTGACAGACGCGCCATGCGGACGCGGACCGGTCACGAGTTGTTGTTCCGCG 315
364 TACATCTGTGATGCGGACGACGACGATCAAGCTCTGAAGAGAGCTGTGTGTGCG 423
316 CAGTGGCGCCCGCGGCGGCTGCTGCCAACGCCATCGGCGCGT-----GAAG 363
424 GAGTGTCTTGGCCCTGCGACAGTGTCTCTAACTGAGTTGAGAGGCTATGGAACAAAGTAC 483
364 TGGTGGCGCCCGGACGAGACCGGATTTCCGCTGATCCCGATGCGTACCCGCGCAGCGG 423
484 TGGAGACGAGAGGCTTCCAGAGGTGCGGAGTGTGTAATGACAAACCCGTAACCGAGAGA 543
424 GTGACAGCTGTGTGCGCGGCGGCGGCGGCGGCTCGCGGACGAGGTGCGTGTGCGCC 483
544 ATCCAGCTGAGTGCACAAAGTGCACG---ACACGCACTTACAAATTCACAGTAGTACT 600
484 TGTGCAAGTGCACGCGCTCACCGCTTCCACCAACAGTGCAGAGTCAAGAGACTTTCGCG 543
601 GCCTGCAAGTGCACGAGTACACCGCGACGACCAAGAGTCCAGTCAAACTTTGAGAGC 660
544 CCGGAGACCGCGCGCGGAGAGAGGTTCGCAAGCCCGCGCGCGCGCGGAGGAGCA 601

```

```

Db      661 ATGTACCTGCAAGCAGTCCAGATCAGAGAGCGGAGAGAGCCAGCAATCCA 718

RESULT 15
US-09-602-877A-92/c
; Sequence 92, Application US/09602877A
; Patent No. 6432707
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Xu, Jianshun
; APPLICANT: Dillon, David C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.446C5
; CURRENT APPLICATION NUMBER: US/09/602,877A
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 92
; LENGTH: 1692
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-602-877A-92

```

```

Query Match      10.3%; Score 65.6; DB 4; Length 1692;
Best Local Similarity 50.7%; Pred. No. 1.6e-06;
Matches 273; Conservative 0; Mismatches 244; Indels 21; Gaps 4;

82 GCCTTCAGAGATGATGCGACAGAGGTGATCCCGGCTTGAGAGTACCCCGAGC---CT 138
1574 GCTTTTAAATGATGCGACAGAAATCTTTATTCATGTGTTAACTGTTCCAGCA 1515
139 CTCTCTGAGAACCAACGACCATGAACCGGCGGAGAGATGAGGACCACTCCC---CAC 195
1514 CACCCGACGACGACGACGCTTGAATCAAGCAGAAATGAGGACGACATTTCACTAAC 1455
196 CATCCCTATGACCCCAAGATGTGTCCGAGTACAGCTGCCGCGGAGCTGCATACCCGC 255
1454 ACTGAGCTGATGGAACACTGCGGTTCAAGTGGTTGCCGGAATGCGTTCCACCAA 1395
256 TTCTGACAGACGCGCCATGCGGACGCGGACCGGTCACGAGTTGTTGTTCCGCG 315
1394 TACATCTGTGATGCGGACGACGACGATCAGCCCTTGAAGAGAGTGTGTGTGCG 1335
316 CAGTGGCGCCCGCGGCGGCTGCTGCCAACGCCATCGGCGCGCT-----GAAG 363
1334 GAGTGTCTTGGCCCTGCGACAGTGTCTCTAACTGAGTTGAGAGGCTATGGAACAAAGTAC 1275
364 TGGTGGCGCCCGGACGAGACCGGATTTCCGCTGATCCCGATGCGTACCGCGCGAGCGG 423
1274 TGGAGACGAGAGAGTCCCGAGAGTGTGTGTAATGACAAACCCGTAACCGAGAGA 1215
424 GTGACAGCTGTGTGCGCGGCGGCGGCGGCGGCTCGCGCAAGTGTGTTGCGCC 483
1214 ATCCAGCTGACAGTCCCAAGATGCGAGC---ACACGCACTTACAAATTCACAGTAGTACT 1158
484 TGTGCAAGTGCACGCGCTCACCGCTTCCACCAACAGTGCAGAGTCAAGAGACTTTCGCG 543
1157 GCCTGCAAGTGCACGAGTACACCGCGACGACCAAGAGTCCAGTCAAACTTTGAGAGC 1098
544 CCGGAGACCGCGCGCGGAGAGAGGTTCGCAAGCCCGCGCGCGCGCGGAGGAGCA 601
1097 ATGTACCTGCAAGGACGTCAGCATCAAGAGAGCGGAGAGAGCCAGCAAAATCCA 1040

```

Search completed: March 29, 2003, 01:31:02
Job time : 59.1204 secs

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 28, 2003, 14:14:27 ; Search time 11.1467 Seconds
(without alignments)
479.722 Million cell updates/sec

Title: US-09-867-274-4

Perfect score: 1032

Sequence: 1 QGMQAFRNDATEVIGLGEY.....KGRKPRGAKNQAULENAY 185

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 2942292 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfillseq1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1020.5	98.9	211	4	US-09-449-218D-12	Sequence 12, Appl
2	982.5	95.2	213	4	US-09-449-218D-14	Sequence 14, Appl
3	951.5	92.2	213	4	US-09-449-218D-10	Sequence 10, Appl
4	932.5	90.4	213	4	US-09-449-218D-2	Sequence 2, Appl
5	923.5	89.5	213	4	US-09-449-218D-6	Sequence 6, Appl
6	923.5	89.5	213	4	US-09-449-218D-8	Sequence 8, Appl
7	839	81.3	176	4	US-09-449-218D-16	Sequence 16, Appl
8	330	32.0	206	4	US-09-188-930-159	Sequence 159, Appl
9	330	32.0	206	4	US-09-188-930-286	Sequence 286, App
10	328	31.8	206	1	US-08-468-847B-2	Sequence 2, Appl
11	328	31.8	206	1	US-08-468-847B-20	Sequence 20, Appl
12	90.5	8.8	184	4	US-09-040-229B-6	Sequence 6, Appl
13	88.5	8.6	184	4	US-09-040-229B-4	Sequence 4, Appl
14	88.5	8.6	184	4	US-09-449-218D-42	Sequence 42, Appl
15	88.5	8.6	184	4	US-09-040-229B-2	Sequence 2, Appl
16	86	8.3	184	4	US-09-040-229B-9	Sequence 9, Appl
17	84	8.1	830	4	US-09-562-737-35	Sequence 35, Appl
18	83.5	8.1	246	4	US-09-336-093-5	Sequence 5, Appl
19	82	7.9	301	3	US-08-303-861-21	Sequence 21, Appl
20	82	7.9	301	4	US-09-011-073A-1	Sequence 1, Appl
21	82	7.9	301	4	US-09-230-421-2	Sequence 2, Appl
22	80.5	7.8	270	4	US-08-878-474-1	Sequence 1, Appl
23	79.5	7.7	735	3	US-09-191-647-9	Sequence 9, Appl
24	79.5	7.7	735	4	US-09-540-245A-9	Sequence 9, Appl
25	79.5	7.7	735	4	US-09-540-153-9	Sequence 9, Appl
26	79	7.7	1213	4	US-09-413-814-79	Sequence 79, Appl
27	77.5	7.5	188	1	US-08-469-427A-5	Sequence 5, Appl

28	77.5	7.5	188	2	US-08-609-443B-5	Sequence 5, Appl
29	77.5	7.5	188	2	US-08-569-063C-5	Sequence 5, Appl
30	77.5	7.5	188	4	US-08-851-896-5	Sequence 5, Appl
31	77.5	7.5	195	1	US-08-469-427A-7	Sequence 7, Appl
32	77.5	7.5	195	2	US-08-609-443B-7	Sequence 7, Appl
33	77.5	7.5	195	2	US-08-569-063C-7	Sequence 7, Appl
34	77.5	7.5	195	4	US-08-851-896-7	Sequence 7, Appl
35	77	7.5	55	1	US-08-469-427A-3	Sequence 3, Appl
36	77	7.5	55	2	US-08-609-443B-3	Sequence 3, Appl
37	77	7.5	55	2	US-08-569-063C-3	Sequence 3, Appl
38	77	7.5	55	4	US-08-851-896-3	Sequence 3, Appl
39	77	7.5	470	2	US-08-946-241B-2	Sequence 2, Appl
40	77	7.5	470	3	US-09-309-053-2	Sequence 2, Appl
41	77	7.5	479	2	US-08-946-241B-9	Sequence 9, Appl
42	77	7.5	479	3	US-09-309-053-9	Sequence 9, Appl
43	76.5	7.4	145	1	US-08-298-189B-1	Sequence 1, Appl
44	76.5	7.4	145	1	US-08-475-213-10	Sequence 10, Appl
45	76.5	7.4	145	2	US-08-395-23B-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-449-218D-12
; Sequence 12, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepers, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-449-218D-12
Query Match 98.9%; Score 1020.5; DB 4; Length 211;
Best Local Similarity 98.4%; Pred. No. 1.2e-98;
Matches 185; Conservative 0; Mismatches 0; Indels 3; Gaps 1;
QY 1 QGMQAFRNDATEVIGLGEYPPPPNNQNRANGRRPPHPPYDAIVDSYSCRELIHY 60
DB 24 QGMQAFRNDATEVIGLGEYPPPPNNQNRANGRRPPHPPYDAIVDSYSCRELIHY 83
QY 61 TRFLTDGPRSKAPVTELYSCGCGPARLLPVAIGRVKWRNPGDPFCIPRYAORVQ 120
DB 84 TRFLTDGPRSKAPVTELYSCGCGPARLLPVAIGRVKWRNPGDPFCIPRYAORVQ 143
QY 121 LILPGGAARSRKRVLVASCCKRLTRFPHNSELKDQFGETARPOKGRKPRP--GAKAN 177
DB 144 LILPGGAARSRKRVLVASCCKRLTRFPHNSELKDQFGETARPOKGRKPRGAGAKAN 203
QY 178 QALEENAY 185
DB 204 QALEENAY 211
RESULT 2
US-09-449-218D-14
; Sequence 14, Application US/09449218D
; Patent No. 6395511


```

; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepert, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-449-218D-14

```

```

Query Match          95.2%; Score 982.5; DB 4; Length 213;
Best Local Similarity 94.2%; Pred. No. 1.1e-94;
Matches 179; Conservative 3; Mismatches 3; Indels 5; Gaps 2;

```

```

QY 1 QGMOAFKNDATFVIGLGEYEPPEPP--ENNQTNRAENGGRPHHPYDAKDVSEYSCREL 58
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
24 QGMOAFKNDATFVIGLGEYEPPEPP--ENNQTNRAENGGRPHHPYDAKDVSEYSCREL 83
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 59 HYTRFLTDGPCRSKAPVTELVCSGCCGPARLLPNAIGRVKWRPNRNGDFRCIDPRYRAQR 118
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
84 HYTRFVTDGPCRSKAPVTELVCSGCCGPARLLPNAIGRVKWRPNRNGDFRCIDPRYRAQR 143
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 119 VOLLCPGGAAPRSKRVRLVASCKCKRLTRFNOSLKDGFETARPQGRKRRP---GAK 175
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 144 VOLLCPGGAAPRSKRVRLVASCKCKRLTRFNOSLKDGFETARPQGRKRRP---GAK 203
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 176 ANQAELENNAY 185
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 204 ANQAELENNAY 213
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

RESULT 3
US-09-449-218D-10
; Sequence 10, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepert, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Cercopithecus pygerythrus
US-09-449-218D-10

```

```

Query Match          92.2%; Score 951.5; DB 4; Length 213;
Best Local Similarity 90.0%; Pred. No. 1.8e-91;
Matches 171; Conservative 10; Mismatches 4; Indels 5; Gaps 2;
QY 1 QGMOAFKNDATFVIGLGEYEPPEPP--ENNQTNRAENGGRPHHPYDAKDVSEYSCREL 58
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

DB 24 QGMOAFKNDATFVIGLGEYEPPEPP--ENNQTNRAENGGRPHHPYDAKDVSEYSCREL 83
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 59 HYTRFLTDGPCRSKAPVTELVCSGCCGPARLLPNAIGRVKWRPNRNGDFRCIDPRYRAQR 118
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 84 HYTRFVTDGPCRSKAPVTELVCSGCCGPARLLPNAIGRVKWRPNRNGDFRCIDPRYRAQR 143
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 119 VOLLCPGGAAPRSKRVRLVASCKCKRLTRFNOSLKDGFETARPQGRKRRP---GAK 175
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 144 VOLLCPGGAAPRSKRVRLVASCKCKRLTRFNOSLKDGFETARPQGRKRRP---GAK 203
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 176 ANQAELENNAY 185
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 204 ANQAELENNAY 213
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

RESULT 4
US-09-449-218D-2
; Sequence 2, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepert, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: 240083.508
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-449-218D-2

```

```

Query Match          90.4%; Score 932.5; DB 4; Length 213;
Best Local Similarity 88.4%; Pred. No. 1.7e-89;
Matches 168; Conservative 10; Mismatches 8; Indels 5; Gaps 4;
QY 1 QGMOAFKNDATFVIGLGEYEPPEPP--ENNQTNRAENGGRPHHPYDAKDVSEYSCREL 58
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
24 QGMOAFKNDATFVIGLGEYEPPEPP--ENNQTNRAENGGRPHHPYDAKDVSEYSCREL 83
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 59 HYTRFLTDGPCRSKAPVTELVCSGCCGPARLLPNAIGRVKWRPNRNGDFRCIDPRYRAQR 118
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 84 HYTRFVTDGPCRSKAPVTELVCSGCCGPARLLPNAIGRVKWRPNRNGDFRCIDPRYRAQR 143
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 119 VOLLCPGGAAPRSKRVRLVASCKCKRLTRFNOSLKDGFETARPQGRKRRP---GAK 175
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 144 VOLLCPGGAAPRSKRVRLVASCKCKRLTRFNOSLKDGFETARPQGRKRRP---GAK 203
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 176 ANQAELENNAY 185
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 204 ANQAELENNAY 213
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

RESULT 5
US-09-449-218D-6
; Sequence 6, Application US/09449218D
; Patent No. 6395511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepert, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
US-09-449-218D-6

```


TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
FILE REFERENCE: 240083.508
CURRENT FILING DATE: 1999-11-24
NUMBER OF SEQ ID NOS: 45
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 213
TYPE: PRT
ORGANISM: Homo sapien
US-09-449-218D-6

Query Match 89.5%; Score 923.5; DB 4; Length 213;
Best Local Similarity 87.9%; Pred. No. 1.5e-88;
Matches 167; Conservative 10; Mismatches 8; Indels 5; Gaps 2;

1 QGWAFFNDATFVIGLGEYEPPE--ENNQTNRANENGRPHHPYDAKDVSEYSCREL 58
24 QGWAFFNDATFVIGLGEYEPPELENNKTNRANENGRPHHPETKDVSEYSCREL 83
59 HYTFPLTDGPRSAKPYTELVCSCGCCPARLLPNAIGRVKWRNGDFRCIPRYAOR 118
84 HFTYVTDGPRSAKPYTELVCSCGCCPARLLPNAIGRVKWRNGDFRCIPRYAOR 143
119 VOLLCPGGAAPRSKRVLVASCCKRLTRFHNOSLKDGPETAPKGRKRP--GAK 175
144 VOLLCPGGAAPRSKRVLVASCCKRLTRFHNOSLKDGPETAPKGRKRP--GAK 175
176 ANQAELENAY 185
204 ANQAELENAY 213

RESULT 6
US-09-449-218D-8
Sequence 8, Application US/09449218D
Patent No. 6395511
GENERAL INFORMATION:
APPLICANT: Brunkow, Mary E.
APPLICANT: Galas, David J.
APPLICANT: Kovacevich, Brian
APPLICANT: Mulligan, John T.
APPLICANT: Paepker, Bryan W.
APPLICANT: Van Ness, Jeffrey
APPLICANT: Winkler, David G.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
FILE REFERENCE: 240083.508
CURRENT FILING DATE: 1999-11-24
NUMBER OF SEQ ID NOS: 45
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 8
LENGTH: 213
TYPE: PRT
ORGANISM: Homo sapien
US-09-449-218D-8

Query Match 89.5%; Score 923.5; DB 4; Length 213;
Best Local Similarity 87.9%; Pred. No. 1.5e-88;
Matches 167; Conservative 10; Mismatches 8; Indels 5; Gaps 2;
1 QGWAFFNDATFVIGLGEYEPPE--ENNQTNRANENGRPHHPYDAKDVSEYSCREL 58
24 QGWAFFNDATFVIGLGEYEPPELENNKTNRANENGRPHHPETKDVSEYSCREL 83
59 HYTFPLTDGPRSAKPYTELVCSCGCCPARLLPNAIGRVKWRNGDFRCIPRYAOR 118
84 HFTYVTDGPRSAKPYTELVCSCGCCPARLLPNAIGRVKWRNGDFRCIPRYAOR 143
119 VOLLCPGGAAPRSKRVLVASCCKRLTRFHNOSLKDGPETAPKGRKRP--GAK 175

DB 144 VOLLCPGGAAPRSKRVLVASCCKRLTRFHNOSLKDGPETAPKGRKRP--GAK 203
QY 176 ANQAELENAY 185
DB 204 ANQAELENAY 213

RESULT 7
US-09-449-218D-16
Sequence 16, Application US/09449218D
Patent No. 6395511
GENERAL INFORMATION:
APPLICANT: Brunkow, Mary E.
APPLICANT: Galas, David J.
APPLICANT: Kovacevich, Brian
APPLICANT: Mulligan, John T.
APPLICANT: Paepker, Bryan W.
APPLICANT: Van Ness, Jeffrey
APPLICANT: Winkler, David G.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
FILE REFERENCE: 240083.508
CURRENT FILING DATE: 1999-11-24
NUMBER OF SEQ ID NOS: 45
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 16
LENGTH: 176
TYPE: PRT
ORGANISM: Bos taurus
US-09-449-218D-16

Query Match 81.3%; Score 839; DB 4; Length 176;
Best Local Similarity 86.4%; Pred. No. 7.2e-80;
Matches 152; Conservative 11; Mismatches 9; Indels 4; Gaps 2;

8 NDATFVIGLGEYEPPE--ENNQTNRANENGRPHHPYDAKDVSEYSCRELHYTRFLTD 66
1 NDATFVIGLGEYEPPELENNKTNRANENGRPHHPETKDVSEYSCRELHFTYVTD 60
67 GPCRSKPYTELVCSCGCCPARLLPNAIGRVKWRNGDFRCIPRYAORVOLLCPG 126
61 GPCRSKPYTELVCSCGCCPARLLPNAIGRVKWRNGDFRCIPRYAORVOLLCPG 120
127 AARSRKRVLVASCCKRLTRFHNOSLKDGPETAPKGRKRP--GAK 179
121 AARSRKRVLVASCCKRLTRFHNOSLKDGPETAPKGRKRP--GAK 176

RESULT 8
US-09-188-930-159
Sequence 159, Application US/09188930A
Patent No. 6150502
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Onuet, Rene
APPLICANT: Marison, James Greg
TITLE OF INVENTION: Compositions Isolated From Skin Cells
FILE REFERENCE: 11000.1011c1
CURRENT FILING DATE: 1998-11-09
NUMBER OF SEQ ID NOS: 348
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 159
LENGTH: 206
TYPE: PRT
ORGANISM: mouse
US-09-188-930-159

Query Match 32.0%; Score 330; DB 4; Length 206;

MEDIUM TYPE: 3.5 INCH DISKETTE
COMPUTER: IBM PS/2
OPERATING SYSTEM: MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,847B
FILING DATE: 6 June 1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: MULLINS, J.G.
REGISTRATION NUMBER: 33,073
REFERENCE/DOCKET NUMBER: 325800-442
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-994-1700
TELEFAX: 201-994-1744
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 206 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS:
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
US-08-468-847B-20

Query Match 31.8%; Score 328; DB 1; Length 206;
Best Local Similarity 39.9%; Pred. No. 1,4e-26;
Matches 73; Conservative 36; Mismatches 62; Indels 12; Gaps 7;

QY 5 AFRNDATEVPGGEYEP-PPENNNTMRAENGRRPHHRYDAKVS---EYSCRELYH 60
DB 23 AFRNDATEVPGGEYEP-PPENNNTMRAENGRRPHHRYDAKVS---EYSCRELYH 80
QY 61 TRFLTPDPCSAKVTETLVSGGCGPARLLPNAIG---RYKWM-RPNDPFCRIPDEYRA 116
DB 81 TRFLTPDPCSAKVTETLVSGGCGPARLLPNAIG---RYKWM-RPNDPFCRIPDEYRA 140
QY 117 QRVOLLPFGAAPRSKRVLVASCKGRLTRFNQSELKPGFETAP-QKGRKPRGAK 175
DB 141 QRVOLLPFGAAPRSKRVLVASCKGRLTRFNQSELKPGFETAP-QKGRKPRGAK 199
QY 176 ANQ 178
DB 200 SSK 202

RESULT 12
US-09-040-229B-6
Sequence 6, Application US/09040229B
Patent No. 6432410
GENERAL INFORMATION:
APPLICANT: Hatland, Richard
Hsu, David
TITLE OF INVENTION: Morphogenic Proteins
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/040,229B
FILING DATE: 13-Mar-1998
CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B97-020-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 184 amino acids
TYPE: amino acid
TOPOLOGY: LINEAR
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-040-229B-6

Query Match 8.8%; Score 90.5; DB 4; Length 184;
Best Local Similarity 22.9%; Pred. No. 0.063;
Matches 35; Conservative 23; Mismatches 60; Indels 35; Gaps 8;

QY 18 GEYEP-PPENNNTMRAENGRRPHH---PYDAKDYEVSCRELYHTR--RFLTDG 67
DB 33 GAIPDPDQPDNDSEOMOTQOQSGSRHREKGTSMPEBEVLESQALHTERKYLKD 92
QY 68 PCR-----SAKPVTELVSQGGCGPARLLPNAIGRVKWMRPNDPFCRIPDEYRA 110
DB 93 WCKTQPLKQTHBEGNSGTTINRFQYGCN-SFYIPRHV-----RKBECSFGSCFCK 145
QY 111 PDYRAQRVOLLPFGAAPRSKRVLVASCKGRLTRFNQSELKPGFETAP-QKGRKPRGAK 142
DB 146 PKFTTMTVTLNCPLEQPRKKRITRVYECRC 178

RESULT 13
US-09-040-229B-4
Sequence 4, Application US/09040229B
Patent No. 6432410
GENERAL INFORMATION:
APPLICANT: Hatland, Richard
Hsu, David
TITLE OF INVENTION: Morphogenic Proteins
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/040,229B
FILING DATE: 13-Mar-1998
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B97-020-2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 182 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-040-229B-4

Query Match 8.6%; Score 88.5; DB 4; Length 184;

Best Local Similarity 23.5%; Pred. No. 0.1; Matches 35; Conservative 23; Mismatches 54; Indels 37; Gaps 8;

QY 23 PPEPNNQTMNRAENG-----GRPPHHPYDAKDVSEYSCRELHYT--RFLTDGPCR-70
DB 36 PPDGKGP-NDSEGGQAGPDGRVGRKGKGAALAEVLSSQGEALHTERKYLKRDWCKT 94
QY 71 -----SAKPYTELVCSGGCGPARLLPNAIGRVKWMRPNNGPDR---CIPDRY 114
DB 95 QPLKQTHEDGNSRTIINRFQYGCN-SFYIPRHIRE-----GSFQSCSFCKPKK 147
QY 115 RAQRVOLLCPGGAAP-RSRKRVLVASCKC 142
DB 148 TTMVVTINCPBLQPTTKKRVKQCR 176

RESULT 14

US-09-449-218D-42
; Sequence 42, Application US/09449218D
; Patent No. 639511
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Galas, David J.
; APPLICANT: Kovacevich, Brian
; APPLICANT: Mulligan, John T.
; APPLICANT: Paepert, Bryan W.
; APPLICANT: Van Ness, Jeffrey
; APPLICANT: Winkler, David G.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR INCREASING
; FILE REFERENCE: BONE MINERALIZATION
; CURRENT APPLICATION NUMBER: US/09/449,218D
; CURRENT FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 42
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-449-218D-42

Query Match 8.6%; Score 88.5; DB 4; Length 184;

Best Local Similarity 23.4%; Pred. No. 0.1; Matches 36; Conservative 20; Mismatches 55; Indels 43; Gaps 8;

QY 23 PPEPNNQ-----TMNRAENGRRPHHPYDAKDVSEYSCRELHYT--RFLTD 66
DB 34 PPDKAQHNDSEGTQSPQPGSRNRGRGGRGTAMP--GEEVLSSQGEALHTERKYLKR 91
QY 67 GPCR-----SAKPYTELVCSGGCGPARLLPNAIGRVKWMRPNNGPDR---C 109
DB 92 DWCKTQPLKQTHEDGNSRTIINRFQYGCN-SFYIPRHIRE-----RKEGSGFQSCSF 144
QY 110 IPDRYRAQRVOLLCPGGAAP-RSRKRVLVASCKC 142
DB 145 KPCKFTTMVVTINCPBLQPTTKKRVKQCR 176

RESULT 15

US-09-040-229B-2
; Sequence 2, Application US/09040229B
; Patent No. 6432410
; GENERAL INFORMATION:
; APPLICANT: Harland, Richard
; Hsu, David
; TITLE OF INVENTION: Morphogenic Proteins
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 75 DENISE DRIVE
; CITY: HILLSBOROUGH

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94010

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/040,229B

FILING DATE: 13-Mar-1998

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: OSMAN, RICHARD A

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: B97-020-2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (650) 343-4341

TELEFAX: (650) 343-4342

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 184 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-040-229B-2

Query Match 8.6%; Score 88.5; DB 4; Length 184;

Best Local Similarity 23.4%; Pred. No. 0.1; Matches 36; Conservative 20; Mismatches 55; Indels 43; Gaps 8;

QY 23 PPEPNNQ-----TMNRAENGRRPHHPYDAKDVSEYSCRELHYT--RFLTD 66
DB 34 PPDKAQHNDSEGTQSPQPGSRNRGRGGRGTAMP--GEEVLSSQGEALHTERKYLKR 91
QY 67 GPCR-----SAKPYTELVCSGGCGPARLLPNAIGRVKWMRPNNGPDR---C 109
DB 92 DWCKTQPLKQTHEDGNSRTIINRFQYGCN-SFYIPRHIRE-----RKEGSGFQSCSF 144
QY 110 IPDRYRAQRVOLLCPGGAAP-RSRKRVLVASCKC 142
DB 145 KPCKFTTMVVTINCPBLQPTTKKRVKQCR 176

Search completed: March 28, 2003, 14:20:15
Job time: 11.3467 secs

Amgen, Inc. (US)
 FEATURES Location/Qualifiers
 source 1..759
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 125 a 282 c 244 g 108 t
 ORIGIN
 Query Match 100.0%; Score 759; DB 6; Length 759;
 Best Local Similarity 100.0%; Pred. No. 1.8e-106;
 Matches 759; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 TACTGGAAGTGGCGTCCCTCTCTGCTGATACCATGACGCTCCACTGAGCTGTGT 60
 Db 1 TACTGGAAGTGGCGTCCCTCTCTGCTGATACCATGACGCTCCACTGAGCTGTGT 60
 61 CTGCTCTGCTGCTGCTGATACACAGCTTCCGTGTAGTGAAGGCGGAGGTGACAGGG 120
 Db 61 CTGCTCTGCTGCTGCTGATACACAGCTTCCGTGTAGTGAAGGCGGAGGTGACAGGG 120
 121 TTCAAGATGATGCGACGAGAAATCATCCCGAGCTGGAGATACCCCGAGCTCCACCG 180
 Db 121 TTCAAGATGATGCGACGAGAAATCATCCCGAGCTGGAGATACCCCGAGCTCCACCG 180
 181 GAGCTGAGAACCAACAGACCATGAAACGGGCGGAGAACGAGGCGGCTCCCTCCACAC 240
 Db 181 GAGCTGAGAACCAACAGACCATGAAACGGGCGGAGAACGAGGCGGCTCCCTCCACAC 240
 241 CCTTTGAGACCAAAAGAGTGTCCGAGTACAGCTGCCGAGCTGCACTTCACTCCGCTAC 300
 Db 241 CCTTTGAGACCAAAAGAGTGTCCGAGTACAGCTGCCGAGCTGCACTTCACTCCGCTAC 300
 301 GTGACGATGGGCGGCGGAGCGGCAAGCGGCTCAACGAGCTGTGTCTCCGCGACG 360
 Db 301 GTGACGATGGGCGGCGGAGCGGCAAGCGGCTCAACGAGCTGTGTCTCCGCGACG 360
 361 TGGCGCCGCGCGCGCTGCTGCCCAAGCCATCGGCGCGCAAGTGTGGCACTAGT 420
 Db 361 TGGCGCCGCGCGCGCTGCTGCCCAAGCCATCGGCGCGCAAGTGTGGCACTAGT 420
 421 GGGCGCCGATCTTCGCTGATCCCGACCGCTACCGCGCGGAGCGGTGACGCTGTGT 480
 Db 421 GGGCGCCGATCTTCGCTGATCCCGACCGCTACCGCGCGGAGCGGTGACGCTGTGT 480
 481 CCGGCTGATGAGGCGCGCGCGCGCGCAAGGTGAGCTGTGAGCTCTGATGATGCAAG 540
 Db 481 CCGGCTGATGAGGCGCGCGCGCGCGCAAGGTGAGCTGTGAGCTCTGATGATGCAAG 540
 541 CGCTCAACCGCTTCCACCAACAGTCGAGCTCAAGGACTTCCGGAACCGAGCGCTCG 600
 Db 541 CGCTCAACCGCTTCCACCAACAGTCGAGCTCAAGGACTTCCGGAACCGAGCGCTCG 600
 601 CCGGAGAGGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
 Db 601 CCGGAGAGGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
 661 CTGAGAGACGCTTACTAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
 Db 661 CTGAGAGACGCTTACTAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720
 721 GAACCG 759
 Db 721 GAACCG 759
 RESULT 2
 AF331844 2296 bp mRNA linear pri 06-MAR-2001
 LOCUS Homo sapiens SOST (SOST) mRNA, complete cds.
 DEFINITION AF331844
 ACCESSION AF331844.1 GI:13236417
 VERSION AF331844.1
 KEYWORDS Homo sapiens.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 2296)
 AUTHORS Balcells, W., Ebeling, M., Patel, N., van Hul, E., Olson, P.,
 Dlozezi, M., Laczka, C., Wuyts, W., van den Ende, J., Willems, P.,
 Paes-Alves, A.F., Hill, S., Bueno, M., Ramos, P., J., Tacon, P.,
 Dikkers, F.G., Stratakis, C., Lindpaintner, K., Vickery, B.,
 Foerzler, D. and Van Hul, W.
 TITLE Increased bone density in sclerosteosis is due to the deficiency of
 a novel secreted protein (SOST)
 JOURNAL Hum. Mol. Genet. 10 (5), 537-543 (2001)
 MEDLINE 21096930
 PUBMED 1181578
 REFERENCE 2 (bases 1 to 2296)
 AUTHORS Balcells, W., Ebeling, M., Patel, N., Vickery, B., Foerzler, D. and Van
 Hul, W.
 TITLE Direct Submission
 JOURNAL Submitted (22-DEC-2000) Medical Genetics, University of Antwerp,
 Universiteitplein 1, Antwerp B2610, Belgium
 FEATURES Location/Qualifiers
 source 1..2296
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="17"
 /map="17q12-q21"
 1..2296
 /gene="SOST"
 38..679
 /gene="SOST"
 /note="mutations may cause sclerosteosis"
 /codon_start=1
 /product="SOST"
 /protein_id="AAK16158.1"
 /db_xref="GI:13236418"
 /translation="MOPLALCLVCLVHTAFRVEQSGWQAFKNDATIEIPLEGYP
 EEPLENNKTNRAENGRRPHNPEPTKDVSYSCRELFRTYVTDGCRSAKPYTE
 LVSCGCGPARLLPNAIGRWKRPSPGDPKCPIDRAQROVLLCPGGEAPARAYR
 LVASCKKRLTRHNOSELKDFETBARPQKGRKPRPARSAANQALEBNAY"
 BASE COUNT 576 a 631 c 607 g 482 t
 ORIGIN
 Query Match 100.0%; Score 759; DB 9; Length 2296;
 Best Local Similarity 100.0%; Pred. No. 1.4e-106;
 Matches 759; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 TACTGGAAGTGGCGTCCCTCTCTGCTGATACCATGACGCTCCACTGAGCTGTGT 60
 Db 2 TACTGGAAGTGGCGTCCCTCTCTGCTGATACCATGACGCTCCACTGAGCTGTGT 61
 61 CTGCTCTGCTGCTGCTGATACACAGCTTCCGTGTAGTGAAGGCGGAGGTGACAGGG 120
 Db 62 CTGCTCTGCTGCTGCTGATACACAGCTTCCGTGTAGTGAAGGCGGAGGTGACAGGG 121
 121 TTCAAGATGATGCGACGAGAAATCATCCCGAGCTGGAGATACCCCGAGCTCCACAC 180
 Db 122 TTCAAGATGATGCGACGAGAAATCATCCCGAGCTGGAGATACCCCGAGCTCCACAC 181
 181 GAGCTGAGAACCAACAGACCATGAAACGGGCGGAGAACGAGGCGGCTCCCTCCACAC 240
 Db 182 GAGCTGAGAACCAACAGACCATGAAACGGGCGGAGAACGAGGCGGCTCCCTCCACAC 241
 241 CCTTTGAGACCAAAAGAGTGTCCGAGTACAGCTGCCGAGCTGCACTTCACTCCGCTAC 300
 Db 242 CCTTTGAGACCAAAAGAGTGTCCGAGTACAGCTGCCGAGCTGCACTTCACTCCGCTAC 301
 301 GTGACGATGGGCGGCGGAGCGGCAAGCGGCTCAACGAGCTGTGTCTCCGCGACG 360
 Db 302 GTGACGATGGGCGGCGGAGCGGCAAGCGGCTCAACGAGCTGTGTCTCCGCGACG 361
 361 TGGCGCCGCGCGCGCTGCTGCCCAAGCCATCGGCGCGCAAGTGTGGCACTAGT 420
 Db 362 TGGCGCCGCGCGCGCTGCTGCCCAAGCCATCGGCGCGCAAGTGTGGCACTAGT 421

[illegible]

ORIGIN

Query Match	100.0%	Score 759	DB 9	Length 2323
Best Local Similarity	100.0%	Pred. No. 1.4e-106		
Matches 759	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY 1	TACTGGAAGGTGACGTCGTCCTCTCTGTGAGTCAATGCAAGTCTCCACATCTGGCCCTGTGT	60		
pb 12	TACTGGAAGGTGACGTCGTCCTCTCTGTGAGTCAATGCAAGTCTCCACATCTGGCCCTGTGT	71		
QY 61	CTGTGTCGCTGT	120		
Db 72	CTGTGTCGCTGT	131		
QY 121	TTCAGAGTATGATGCAACGAAATCATCCCGAGCTCGAGAGTACCCCGAGCTTCCACCG	180		
Db 132	TTCAGAGTATGATGCAACGAAATCATCCCGAGCTCGAGAGTACCCCGAGCTTCCACCG	191		
QY 181	GAGCTGAGAAACAACAAGACATGAAACCGGGGGGAGAAACGAGGGGGGGCCCTCCACACAC	240		
Db 192	GAGCTGAGAAACAACAAGACATGAAACCGGGGGGAGAAACGAGGGGGGGCCCTCCACACAC	251		
QY 241	CCCTTTAGACCAAAAGACGTGTCCGAGTAAAGCTGCCGAGCTGTCACTTCAACCCGCTAC	300		
Db 252	CCCTTTAGACCAAAAGACGTGTCCGAGTAAAGCTGCCGAGCTGTCACTTCAACCCGCTAC	311		
QY 301	GTAACCGATGAGCCGTGTCGAGTAAAGCTGCCGAGCTGTCACTTCAACCCGCTAC	360		
Db 312	GTAACCGATGAGCCGTGTCGAGTAAAGCTGCCGAGCTGTCACTTCAACCCGCTAC	371		
QY 361	TGCGGCGCGGGGCGCGCTGTCGCGCAAGCATGCGCGCGCAAGTGTGTGTGTGTGTGTGTGT	420		
Db 372	TGCGGCGCGGGGCGCGCTGTCGCGCAAGCATGCGCGCGCAAGTGTGTGTGTGTGTGTGTGT	431		
QY 421	GCGGCGCGACTTCCGCTGATCCCGACCGCTAACCGCGCGACGCTGTGACGTCTGTGT	480		
Db 432	GCGGCGCGACTTCCGCTGATCCCGACCGCTAACCGCGCGACGCTGTGACGTCTGTGT	491		
QY 481	CCCGGTGTGAGGCGCGCGCGCGCGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	540		
Db 492	CCCGGTGTGAGGCGCGCGCGCGCGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	551		
QY 541	GCGCTCAACCGCTTCCACAACAGTCCGAGTCAAGACTTGTGGAGACCGAGCGCTGTGG	600		
Db 552	GCGCTCAACCGCTTCCACAACAGTCCGAGTCAAGACTTGTGGAGACCGAGCGCTGTGG	611		
QY 601	CCGCAAGAGGCGCGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG	660		
Db 612	CCGCAAGAGGCGCGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG	671		
QY 661	CTGGAAGACCGCTTCTAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT	720		
Db 672	CTGGAAGACCGCTTCTAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT	731		
QY 721	GAAACCGGCGCGCAATTTCTGTCTCTGTGCGCGGTGT	759		
Db 732	GAAACCGGCGCGCAATTTCTGTCTCTGTGCGCGGTGT	770		
RESULT 4				
LOCUS	AX056687	2329 bp	DNA	linear
DEFINITION	Sequence 19 from Patent WO0075317.			
ACCESSION	AX056687			
VERSION	AX056687.1	GI:12309667		
KEYWORDS				
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
	1 (bases 1 to 2329)			
	Borstein,D.A., Goddard,A., Gurney,A.L., Smith,V., Matanabe,C.K. and			
	Wood,W.I.			

